Single-Molecule Imaging Reveals that Argonaute Re-Shapes the Properties of its Nucleic Acid Guides: A Dissertation

SINGLE-MOLECULE IMAGING REVEALS THAT ARGONAUTE RE-SHAPES THE PROPERTIES OF ITS NUCLEIC ACID GUIDES

A Dissertation Presented

By

WILLIAM EDWARD SALOMON

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 7, 2015

BIOCHEMISTRY AND MOLECULAR PHARMACOLOGY GRADUATE PROGRAM

SINGLE-MOLECULE IMAGING REVEALS THAT ARGONAUTE RE-SHAPES THE PROPERTIES OF ITS NUCLEIC ACID GUIDES

A Dissertation Presented

By

WILLIAM EDWARD SALOMON

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

_____

Phillip D. Zamore, Ph.D., Thesis Adviser

_____

Anthony Carruthers, Ph.D., Member of Committee

_____

Craig C. Mello, Ph.D., Member of Committee

_____

Sean P. Ryder, Ph.D., Member of Committee

_____

Ian J. MacRae, Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

_____

Thoru Pederson, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school.

_____

Anthony Carruthers, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

Biochemistry and Molecular Pharmacology Program
December 7, 2015

DEDICATION

I dedicate this thesis to my family,

Mom, Dad, Stephen, and Zack

and

Mark A. Taylor,

you left too soon my friend but your positivity

and spirit have always reminded me to live life to the fullest

Acknowledgments

There is an ancient proverb that says, “It takes a village to raise a child.” In the same vein, I say, “It takes a lab for a graduate student to earn a PhD,” and this could not be any truer for my experience in the Zamore Lab at UMass Medical School. First, I want to thank the head of our ‘village’ Phillip Zamore, my thesis advisor and mentor. Phil’s philosophy that, seeing is believing, but *measuring* is *knowing,* is one that will serve me well for the rest of my career. I am grateful to have had this opportunity to learn from Phil on how to become a rigorous biochemist that uses quantitative measurements to understand biology (also known as “the program”).

My colleagues in the Zamore lab are more like friends and it is their friendships that made coming to ‘work’ feel nothing of the sort. My thesis would not have been possible without my buddies in the ‘AGO club’, Jennifer “Jenny” Broderick and Carlos “Fabián” Flores-Jasso whom I have shared countless cups of coffee with and many hours of discussion. Fabián and Jenny were like the older siblings I never had (I am the oldest among my two siblings). Fabián was my post-doc mentor and I learned many things from him. I enjoyed scheming with him ways to increase the yields of RNA after gel purification and fighting the good fight in the gel room. Jenny’s wisdom, creativity, and advice on anything and everything are things that I could not have gone without during my time in the lab.

My thesis project would not have been possible without the great collaboration I had with Victor Serebrov. Victor and I embarked on using single-molecule techniques to study Argonaute proteins and I have learned a lot from him. I hope to continue to collaborate with Victor in the future. Tracy Lincoln, Chris Matranga, Jenny, and Fabián who relentlessly grilled me in my preparation

for the qualifying exam — you guys are the reason I was so well prepared! Sam Jolly, it was great to collaborate in our research projects and become friends in the process. Liang Meng Wee, a former student and a master at the RNAi cleavage assay. I learned from the world's best and I am appreciative of all the help and knowledge you passed down to me. Amena Arif whom I had the chance to mentor and pass the torch of cleavage assays to; I cannot wait to see your future success! To the rest of my labmates, you all have made my time in the lab an experience I will never forget. I would be remiss if I did not thank those that keep our village from burning down, Gwen Farley, Alicia Boucher, and Tiffanie Gardner — it cannot be said enough, Thank you!

To my friends and colleagues outside the lab at UMass, you make this place special. Outside of school you kept me sane in this sometimes crazy world we call research — whether it was playing softball or volleyball, catching a Sharks or Bravehearts game, or just having a beer at the Abbey, I enjoyed our time together. To my college buddies in the area, Cotton, Courtney, JB, Tone, Walka, just to name a few, you have no idea what it meant to me to just unplug and get away for a little while. Playing some golf, watching the Pats, or standing on ice drinking some beers (ice fishing) — it was just the escape I needed!

Lastly, but in no way the least, is my family and the tremendous support they have given me. My Mom and Dad for their constant encouragement and advice. I don't say it enough, but, I love you and earning a PhD would not have been possible without you both. My two younger brothers, Stephen and Zack, you guys are my best friends. A text from you guys made the hard days in research bearable. You always make me laugh the hardest and I am grateful for the many beers you have bought this poor grad student — now I owe you a few!

— Wes

December, 2015

Abstract

Small RNA silencing pathways regulate development, viral defense, and genomic integrity in all kingdoms of life. An Argonaute (Ago) protein, guided by a tightly bound, small RNA or DNA, lies at the core of these pathways. Argonaute uses its small RNA or DNA to find its target sequences, which it either cleaves or stably binds, acting as a binding scaffold for other proteins. We used Co-localization Single-Molecule Spectroscopy (CoSMoS) to analyze target binding and cleavage by Ago and its guide. We find that both eukaryotic and prokaryotic Argonaute proteins re-shape the fundamental properties of RNA:RNA, RNA:DNA, and DNA:DNA hybridization: a small RNA or DNA bound to Argonaute as a guide no longer follows the well-established rules by which oligonucleotides find, bind, and dissociate from complementary nucleic acid sequences. Counter to the rules of nucleic acid hybridization alone, we find that mouse AGO2 and its guide bind to microRNA targets 17,000 times tighter than the guide without Argonaute. Moreover, AGO2 can distinguish between microRNA-like targets that make seven base pairs with the guide and the products of cleavage, which bind via nine base pairs: AGO2 leaves the cleavage products faster, even though they pair more extensively.

This thesis presents a detailed kinetic interrogation of microRNA and RNA interference pathways. We discovered sub-domains within the previously

defined functional domains created by Argonaute and its bound DNA or RNA guide. These sub-domains have features that no longer conform to the well-established properties of unbound oligonucleotides. It is by re-writing the rules for nucleic acid hybridization that Argonautes allow oligonucleotides to serve as specificity determinants with thermodynamic and kinetic properties more typical of RNA-binding proteins than that of RNA or DNA. Taken altogether, these studies further our understanding about the biology of small RNA silencing pathways and may serve to guide future work related to all RNA-guided endonucleases.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

Copyright Information

ABBREVIATIONS

ΔG: Change in Gibbs free energy
Ago: Argonaute
ABD: Argonaute binding domain
Alexa: Proprietary name of fluorescent dye originally named after the son (Alex) of the founders of Molecular Probes (acquired by Life Technologies)
ATP: Adenosine triphosphate, single letter abbreviation A
C3PO: Component 3 promoter of RISC
C. elegans: *Caenorhabditis elegans*
CAF1: CCR4 associated factor 1
CCR4: Chemokine receptor 4
CTP: Cytidine triphosphate, single letter abbreviation C
DGCR8: DiGeorge syndrome critical region 8
D. melanogaster: *Drosophila melanogaster* (fruit fly)
dCTP: Deoxycytidine triphosphate (dC)
DNA: Deoxyribonucleic acid
DTT: Dithiothreitol
dUTP: deoxyuridine triphosphate, (dU)
E. coli: *Escherichia coli*
endo-siRNA: endogenous siRNA
GDP: Guanosine diphosphate
GTP: Guanosine triphosphate, single letter abbreviation G
GW182: Glycine-tryptophan repeat protein, ~182 kDa
HeLa: Immortal, tumor-derived cell line from the cancer patient Henrietta Lacks. The acquisition of these cells was controversial (1951 at John Hopkins by George Gey) but it is one of the oldest and widely used human cell lines.
HEN1: Hua enhancer 1
HSC: Heat shock cognate protein
HSP: Heat shock protein
k_{cat}: Michaelis-Menten turnover rate
K_D: Dissociation constant
K_M: Michaelis-Menten constant
k_{off}: Dissociation rate constant
k_{on}: Association rate constant
Klenow: Large fragment from *E. coli* DNA polymerase I
La: Sjögren's syndrome type B autoantigen, a RNA-binding protein
let-7: lethal-7, from *C. elegans*

lin-4: cell lineage miRNA, one of the first discovered miRNAs
m^7G: 7-methyl guanosine cap
MEF: Mouse embryonic fibroblast cells
mRNA: Messenger RNA
miRNA: microRNA
MIWI/MILI: Mouse PIWI protein
MOV10L1: Moloney Leukemia Virus 10-Like Protein
NEM: *N*-ethylmaleimide
NOT: General negative regulator of transcription
Oligo: Oligonucleotide, typically a synthetically synthesized DNA or RNA
PABP: Poly(A) binding protein
PAIP: PABP interacting protein
PAN: Poly(A) Specific ribonuclease subunit
PAZ: PIWI/Argonaute/Zwille
PEG: Polyethylene glycol
PCR: Polymerase chain reaction
PIWI: P-element induced wimpy testes
PLD6: phospholipase/endonuclease, also known as mouse Zuc
Pri-miRNA: Primary miRNA
Pre-miRNA: Precursor miRNA
Pol II: RNA polymerase II
Ran-GTP/GDP: Ras-related Nuclear protein bound to GTP or GDP
R2D2: Two dsRNA binding domain (R2) and Dicer-2 associated (D2)
RISC: RNA-induced silencing complex
RNA: Ribonucleic acid
RNAi: RNA interference
RRP44: Ribosomal RNA-processing protein 44 (3′ → 5′ exoribonuclease)
S100: Soluble fraction after centrifugation at 100,000 × gravity, “S100 extract”
SDS: Sodium dodecyl sulfate
siRNA: Small interfering RNA
T7: T-odd, lytic bacteriophage that infects *E. coli* (T7 RNA polymerase source)
TIRF: Total internal reflection fluorescence
TRBP: HIV trans-activating response (TAR) RNA-binding protein
TtAGO: Argonaute protein of *Thermus thermophilus*
UTP: Uridine triphosphate, single letter abbreviation U
V_0: Initial velocity
XRN-1: Exoribonuclease-1 (the major 5′ → 3′ exoribonuclease in cells)
Zuc: Zucchini, germline mutant in *Drosophila melanogaster*

Published Works

The following publications appear in whole or part in the thesis:

Flores-Jasso, C.F., Salomon, W.E., and Zamore, P.D. (2013). Rapid and specific purification of Argonaute-small RNA complexes from crude cell lysates. RNA 19, 271–279.

Salomon, W.E., Jolly, S.M., Moore, M.J., Zamore, P.D., and Serebrov, V. (2015). Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. Cell 162, 84–95.

The following publications were contributed works that are not included in the thesis:

Wee, L.M., Flores-Jasso, C.F., Salomon, W.E., and Zamore, P.D. (2012). Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. Cell 151, 1055–1067.

Broderick, J. A., Salomon, W.E., Ryder, S.P., Aronin, N., and Zamore, P.D. (2011). Argonaute protein identity and pairing geometry determine cooperativity in mammalian RNA silencing. RNA 17, 1858–1869.

Tesz, G.J., Aouadi, M., Prot, M., Nicoloro, S.M., Boutet, E., Amano, S.U., Goller, A., Wang, M., Guo, C.-A., Salomon, W.E., Virbasius, J. V., Baum, R.A., O'Connor, M.J., Soto, E., Ostroff, G.R., Czech, M.P. (2011). Glucan particles for selective delivery of siRNA to phagocytic cells in mice. Biochem. J. 436, 351–362.

CHAPTER I: INTRODUCTION

Small Regulatory RNAs: The discovery of something 'small' that is BIG for Biology

Small Regulatory RNA silencing pathways are present in all kingdoms of life and are involved with cellular processes of development, post-transcriptional regulation, viral defense, and genome integrity (Ameres and Zamore, 2013; Swarts et al., 2014b). The discovery of double-stranded RNA triggering a gene silencing effect is attributed to the seminal work of Fire and Mello where they showed that injecting long double-stranded RNA into the nematode *Caenorhabditis elegans* (*C. elegans*) led to silencing of genes that bear sequence similarity to the injected RNA (Fire et al., 1998). This method of gene silencing was coined as RNA interference (RNAi) and it explained previously reported phenomena in plants, fungi, and worms where exogenously introduced transgenes led to repression of genes with similarity and the observation when either antisense or sense RNA was injected into worm it led to repression (Napoli et al., 1990; van der Krol et al., 1990; Matzke and Matzke, 1995; Powell-Coffman et al., 1996; Rocheleau et al., 1997).

Just prior to the discovery that dsRNA was realized to silence gene expression, it converged with another emerging area of *C. elegans* biology. Small endogenous RNAs, (called at the time small temporal RNAs (stRNAs), *lin-4* and *let-7* had been shown to temporally control developmental timing and became intertwined with RNAi (Lee et al., 1993; Wightman et al., 1993; Reinhart

et al., 2000). The importance of these stRNAs soon was realized as something not just restricted to nematodes but also having conservation in multiple organisms, including plants and mammals (Hamilton and Baulcombe, 1999; Pasquinelli et al., 2000). Concordant with the identification of stRNA sequences, there was a flurry of research to understand the biogenesis of the small RNAs in both the stRNA and RNAi pathways. The convergence of stRNA and RNAi occurred with the finding that they shared a common enzyme called Dicer; this enzyme is responsible for producing small interfering RNA (siRNA) in the RNAi pathway that induced gene silencing or was responsible for the stRNA to control gene expression (Fagard et al., 2000; Pasquinelli et al., 2000; Yang et al., 2000; Ambros, 2001; Bernstein et al., 2001; Hutvágner et al., 2001). To reign in the identification of small regulatory RNAs, a specific set of guidelines was devised to distinguish between endogenous stRNAs and siRNAs that were derived from exogenous sources (*e.g.,* viral long dsRNA); the stRNAs were renamed to microRNAs (miRNAs) and they came with specific criteria to validate them experimentally (Ambros et al., 2003). The increase and improvement of sequencing techniques led to the discovery of additional classes of small regulatory RNAs found in several organisms in all kingdoms of life (Bartel, 2004). Fueled by high-throughput sequencing, another class of small RNAs was discovered called endogenous siRNAs (endo-siRNAs) that are found in fungi, protists, plants, and animals – including humans (Yang and Kazazian, 2006;

Czech et al., 2008; Ghildiyal et al., 2008; Okamura et al., 2008a; Tam et al., 2008; Watanabe et al., 2008) (Mochizuki et al., 2002; Volpe et al., 2002; Xie et al., 2004; Chung et al., 2008; Okamura et al., 2008b). Another class of small regulatory RNAs discovered was the germline-specific class that mapped to repeat associated regions of the genome (known as rasiRNAs, now called PIWI-interacting RNAs (piRNAs)); It is believed that these small RNAs help repress transposon activity and protect genome integrity (Aravin et al., 2001; Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006b; Lau et al., 2006; Saito et al., 2006; Vagin et al., 2006; Batista et al., 2008; Das et al., 2008; Grimson et al., 2008). The piRNA and endo-siRNA classes of small RNAs have specialized functions related to regulation of transposons and/or genome integrity in specific cells, whereas miRNAs are believed to regulate protein coding genes in almost all cells (Bartel, 2009; Ameres and Zamore, 2013). No matter the tissue location or cellular process, small RNA regulatory pathways share a common theme — using a small nucleic acid in complex with a protein to drive genomic regulation; this regulation can include temporal regulation of gene expression in animals, genome re-arrangement in protists, post-transcriptional silencing in fungi, defense against genetic information encoded by a virus, and/or protection from selfish genetic elements such as transposons. The level of regulation by these small RNAs is controlled and

carried out by an assembly of specialized proteins that produce and bind the small RNAs in order for regulation of their cellular targets to take place.

Components of Small Regulatory RNA Silencing Pathways in Mammals

The most important part of the small RNA pathway is the small RNA itself that interacts with an endonuclease in the Argonaute superfamily of proteins that regulate cellular targets. All three types of small RNAs — siRNAs, miRNAs, and piRNAs —have respective protein components that can be broken into 3 categories: biogenesis, factors involved with the formation and association with an Argonaute protein, and then the subsequent regulation of the cellular target.

Biogenesis of miRNAs

The transcription of miRNAs, like all endogenous small RNAs, begins as an RNA polymerase II transcript where the small RNA then undergoes further nuclear processing and export to the cytoplasm (Figure 1.1A–C; Lee et al., 2004a; Li et al., 2013a; Kim et al., 2009). The transcription product or primary-miRNA (pri-miRNA) is processed in a stepwise fashion to form a smaller hairpin (Lee et al., 2002). The pri-miRNA transcript containing a large hairpin is recognized by the microprocessor complex that contains the RNase III endonuclease Drosha and its binding partner DGCR8, where it is cropped to a hairpin known as the pre-miRNA (Lee et al., 2003; Gregory et al., 2004; Han et

al., 2004). The pre-miRNA that is formed by the microprocessor complex bears important features like a 5′ phosphate and 2 nt 3′ end overhang; both are crucial for interaction downstream in miRNA biogenesis (Han et al., 2004; Zeng and Cullen, 2004; Zhang et al., 2004; Park et al., 2011). The pre-miRNA associates with the nuclear export protein exportin-5 and it is exported to the cytoplasm (Yi et al., 2003; Bohnsack et al., 2004; Lund et al., 2004). Prior to nuclear export, there are several examples of miRNA specific regulators that control the microprocessor cleavage steps in addition to miRNA specific factors that can associate with the pre-miRNA prior to export (Ha and Kim, 2014). Once the pre-miRNA is exported to the cytoplasm the RNase III endonuclease Dicer along with its dsRNA-binding protein partner TRBP (HIV Transactivating Response RNA-Binding Protein), bind the pre-miRNA hairpin and cleave off the hairpin loop to form a miRNA duplex of 21–23 nt in length (Figure 1.1A; Bernstein et al., 2001; Hutvágner et al., 2001; Haase et al., 2005). This duplex is now able to interact with an Argonaute protein to form RISC, the RNA-induced silencing complex (Figure 1.1A).

Figure 1.1

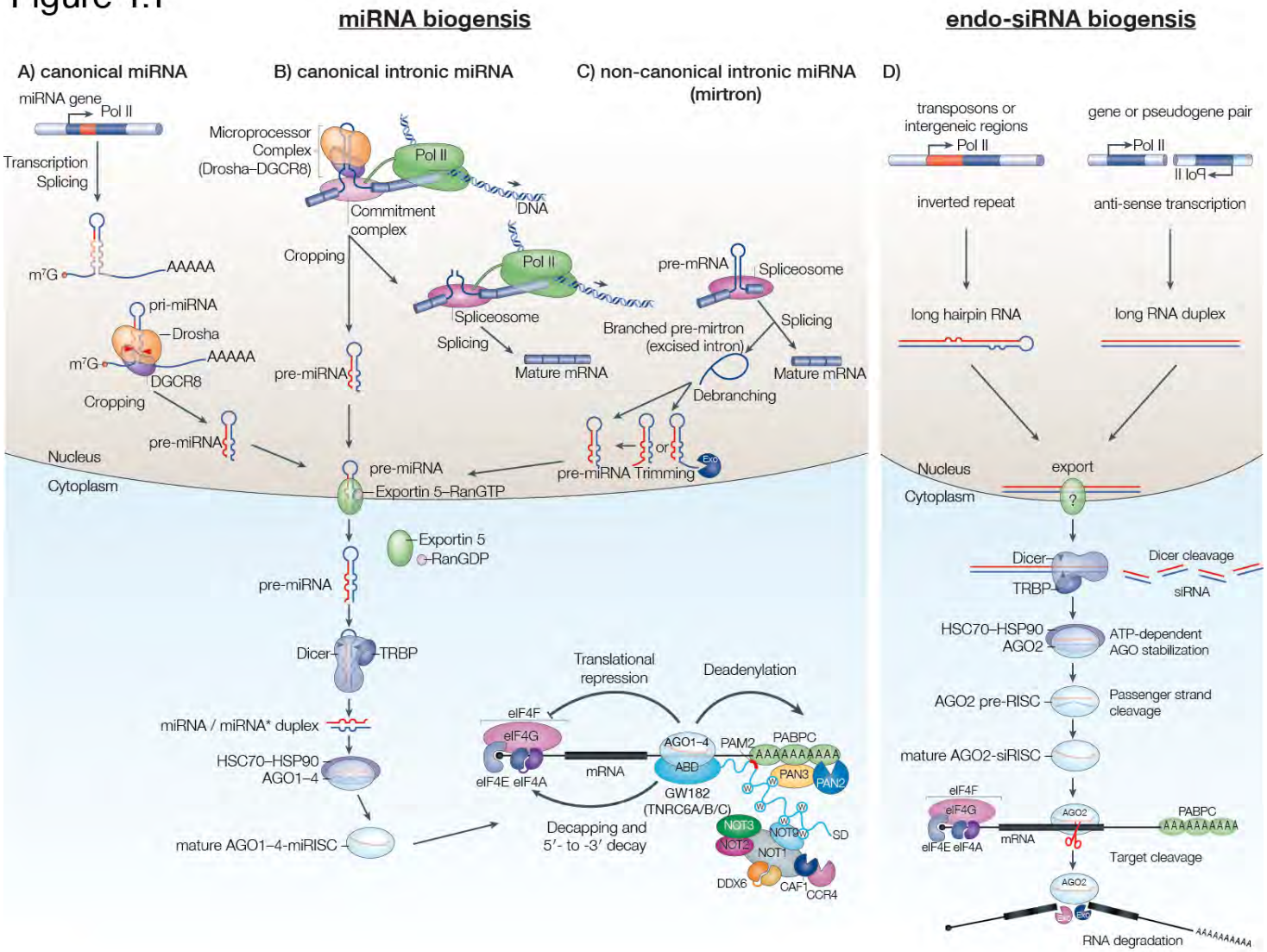

miRNA biogensis
endo-siRNA biogensis
A) canonical miRNA
B) canonical intronic miRNA
C) non-canonical intronic miRNA (mirtron)
D)
miRNA gene
Pol II
Transcription
Splicing
m7G
AAAAA
pri-miRNA
Drosha
DGCR8
Cropping
pre-miRNA
Microprocessor Complex (Drosha–DGCR8)
Commitment complex
DNA
Cropping
Spliceosome
Splicing
Mature mRNA
pre-mRNA
Branched pre-mirtron (excised intron)
Debranching
Nucleus
Cytoplasm
Exportin 5–RanGTP
pre-miRNA Trimming
Exportin 5
RanGDP
Dicer
TRBP
miRNA / miRNA* duplex
HSC70–HSP90
AGO1–4
mature AGO1–4-miRISC
Translational repression
Deadenylation
eIF4F
eIF4G
eIF4E eIF4A
mRNA
AGO1–4
PAM2
PABPC
ABD
PAN3
PAN2
GW182 (TNRC6A/B/C)
NOT3
NOT9
SD
NOT2
NOT1
DDX6
CAF1
CCR4
Decapping and 5′- to -3′ decay
transposons or intergeneic regions
gene or pseudogene pair
inverted repeat
anti-sense transcription
long hairpin RNA
long RNA duplex
export
Dicer cleavage
siRNA
ATP-dependent AGO stabilization
AGO2 pre-RISC
Passenger strand cleavage
mature AGO2-siRISC
AGO2
Target cleavage
RNA degradation

Figure 1.1: miRNA and endogenous siRNA biogenesis, assembly, and regulation pathways in mammals

(A) The biogenesis of miRNAs is primarily transcribed by RNA polymerase II where 50% come from their own gene and typically are part of polycistronic transcription unit. (B) Other miRNAs (~40%) are transcribed from intronic or intergenic locations. (C) A smaller sub-set of miRNAs (~10%) come the excised introns after splicing. After transcription, the pri-miRNAs are processed by the nuclear dsRNA processing complex called the microprocessor complex — Drosha and its binding partner, DGCR8. The newly formed pre-miRNA is exported to the cytoplasm where it can interact with Dicer and its double-stranded binding protein partner, TRBP. The stem-loop of the hairpin is cleaved to form the miRNA/miRNA* duplex. This duplex is assembled into an active AGO protein where the miRNA* strand is ejected and miRNA AGO-RISC is formed. miRNA-RISC partially pairs to its targets where it acts as a scaffold for regulatory factors such as TNRC6A and deadenylation/mRNA regulatory proteins. (D) Endogenous siRNA pathway found in mammals. Endo-siRNAs typically come from transposons, intergeneic regions, or genes/pseudogenes pairs that form a RNA duplex. The long dsRNA is exported to the cytoplasm where it is processed by Dicer-TRBP to form siRNAs. The siRNAs can complex with an active AGO2 protein that will form mature RISC. Endogenous siRNAs that have extensive complementarity to their target can cleave the mRNA thus leaving it susceptible for cellular exonucleases that are commonly found in the exosome (*e.g.,* XRN-1 and RRP44). Figure is adapted with permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology and Nature Review Genetics (Kim et al., 2009; Ha and Kim, 2014; Jonas and Izaurralde, 2015) © 2009–2015.

Biogenesis of siRNAs

The presence of siRNAs in mammals can be broken into two broad categories, those coming from an endogenous or exogenous source. Typically, in the vernacular, siRNAs are referred to as an RNA duplexes ~21 nt in length that are chemically synthesized and introduced exogenously for functional genomic experiments or therapeutic applications (Dorsett and Tuschl, 2004). Endogenous siRNAs (endo-siRNAs) are a specialized class of small RNA found in the germline of mammals, including humans, that are typically derived from transposons that will form dsRNA transcripts (Watanabe et al., 2006; Yang and Kazazian, 2006; Watanabe et al., 2008; Song et al., 2011). The formation of endo-siRNAs is dependent upon the Dicer protein, where loss of Dicer leads to an increase in transposon expression and significant decrease in the levels of endo-siRNAs (Tam et al., 2008; Watanabe et al., 2008). Unlike miRNAs, endo-siRNAs do not get processed in the nucleus by the microprocessor complex, but they do follow the same processing pathway as miRNAs in the cytoplasm — cleavage by Dicer to form an RNA duplex with 5′ phosphate and 2 nt 3′ end overhangs (Kim et al., 2009). The endo-siRNA duplex is then is able to interact with Argonaute to form RISC (Figure 1.1D). For siRNAs that come from exogenous sources there is no requirement for Dicer processing and they can directly associate with Argonaute to go onto forming RISC (Bernstein et al., 2003; Salomon et al., 2010; Betancur and Tomari, 2012).

Biogenesis of piRNAs

The process in which PIWI-interacting small RNAs are made is unlike the Dicer-dependent pathways of miRNAs or endo-siRNAs. The precursor of a piRNA is ssRNA and it is typically only found in the germline of animals (Figure 1.2, model shown in male mouse germline; Vagin et al., 2006; Iwasaki et al., 2015b). In addition to being processed in a Dicer-independent fashion, the length of piRNAs are slightly longer than siRNAs and miRNAs, where they are typically 24–31 nt in length and the 3′ terminal nucleotide bears a 2′-*O*-methyl modification (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006a; Lau et al., 2006; Vagin et al., 2006). The generation of a mature piRNA is not as clear as the miRNA or endo-siRNA pathways due to the lack of good extracts that can recapitulate the pathway in vitro. However, genetic and high-throughput sequencing have informed us a great deal about some of the functions and biogenesis in the pathway. The precursor for a piRNA is long RNA transcript produced by RNA polymerase II and the primary transcript is then fragmented by the phospholipase /endonuclease PLD6 (commonly referred to as Zucchini, name of the fly mutant) where the precursor fragments have enrichment for the 5′ most nucleotide to be a uridine (Girard et al., 2006; Ipsaro et al., 2012; Voigt et al., 2012; Li et al., 2013a; Li et al., 2013b). These fragments are loaded into a PIWI protein, a specialized, germline-specific Argonaute where the 3′ end of the RNA is further trimmed and the terminal nucleotide is modified to a 2′-*O*-methyl

by the ribose methyltransferase HEN1 (Figure 1.2; Vagin et al., 2006; Ohara et al., 2007; Kawaoka et al., 2011; Horwich et al., 2007; Saito et al., 2007; Kamminga et al., 2010). In addition to piRNAs having a distinct biogenesis pathway from miRNAs or endo-siRNAs, they also differ in RISC formation as they have a ssRNA trigger rather than a small RNA duplex. Furthermore, additional piRNAs are made through a secondary piRNA biogenesis pathway that self-amplifies the small RNAs between two different PIWI-Argonaute proteins (Figure 1.2, bottom right). This pathway uses the primary piRNAs to cleave complementary transcripts that then become new piRNAs (secondary piRNAs) that then cleave another primary transcript to form another piRNA with the same sequence as the primary piRNA (Brennecke et al., 2007; Gunawardane et al., 2007; Roovers et al., 2015). Unlike miRNAs and siRNAs, piRNAs biogenesis and assembly with Argonaute proteins are closely related where the mature piRNA is not actually formed until it is bound to a PIWI Argonaute protein (Iwasaki et al., 2015b).

Figure 1.2

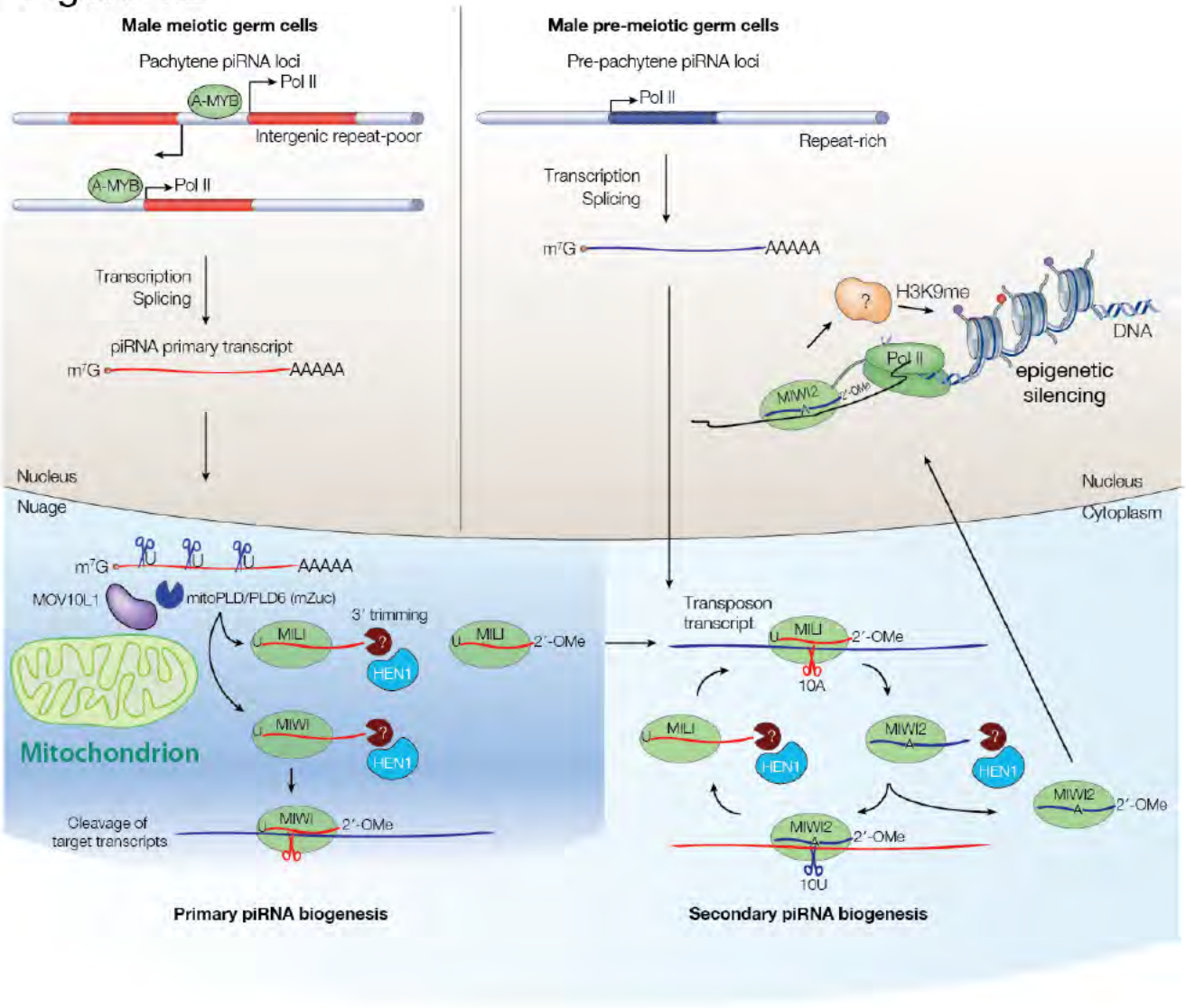

Male meiotic germ cells
Pachytene piRNA loci
Pol II
A-MYB
Intergenic repeat-poor
A-MYB
Pol II
Transcription
Splicing
piRNA primary transcript
m7G
AAAAA
Male pre-meiotic germ cells
Pre-pachytene piRNA loci
Pol II
Repeat-rich
Transcription
Splicing
m7G
AAAAA
?
H3K9me
DNA
Pol II
MIWI2
2′-OMe
epigenetic
silencing
Nucleus
Nuage
Nucleus
Cytoplasm
m7G
AAAAA
MOV10L1
mitoPLD/PLD6 (mZuc)
3′ trimming
MILI
?
HEN1
MILI
2′-OMe
Mitochondrion
MIWI
?
HEN1
Cleavage of
target transcripts
MIWI
2′-OMe
Transposon
transcript
MILI
2′-OMe
10A
MILI
?
HEN1
MIWI2
?
HEN1
MIWI2
2′-OMe
MIWI2
2′-OMe
10U
Primary piRNA biogenesis
Secondary piRNA biogenesis

Figure 1.2: Model of the piRNA pathway in mammals

The piRNA pathway in mammals have a specialized function in the germline for removal mRNA in spermatogenesis, silencing of L1 retrotransposon, and directing CpG DNA methylation on transposons (Aravin et al., 2007; Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008; De Fazio et al., 2011; Reuter et al., 2011). This figure represents a model of the piRNA biogenesis and function in the male germline as it relates to spermatogenesis in the pre-pachytene and pachytene stage. During the pre-pachytene stage, primary piRNAs that were already loaded in MILI can self-amplify to produce secondary piRNAs with the help of MIWI2 in a 'ping-pong'-like amplification loop (right side). The piRNAs in MIWI2 can participate in de novo DNA methylation at non-annotated and transposons regions of the genome. The pachytene class of piRNAs takes place during meiosis and spermatogenesis. The piRNAs are derived from intergeneic regions where their biogenesis is not well understood. The current model has Pol II transcripts (transcription factor A-Myb regulated shown, left side) that are spliced and exported to the specialized compartment of the cytoplasm known as nuage. The transcripts will get processed by PLD6/Zucchini that cleaves the transcript and enriches for 5′ uridine. The fragments are loaded into MIWI where it regulates transcripts via a slicing mechanism to decrease mRNA levels, such as L1 retrotransposon. Figure uses elements with permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology, (Ha and Kim, 2014) © 2014.

Assembly of the siRNA and miRNA RNA-induced Silencing Complexes

The complex formation of a small RNA duplex and the formation of RISC — an Argonaute bound with a single-stranded small RNA — can be broken into two steps, pre-RISC and mature RISC formation (Kawamata and Tomari, 2010; Kobayashi and Tomari, 2015). The formation of pre-RISC and AGO protein association in mammals appears to be the same for both miRNAs and siRNAs, this is in contrast to the fruit fly *Drosophila melanogaster* that is able to discriminate between the two classes of small RNAs (Tomari et al., 2007; Yoda et al., 2010). An example of how flies sort their small RNA is seen in RNAi, where there is a formal RISC loading complex (RLC), made up of Dicer-2 and R2D2 and is a required for RISC formation (Tomari et al., 2007; Czech et al., 2009; Ghildiyal et al., 2010). The siRNA duplex is 'handed off' to fly AGO2 (the RNAi specific AGO protein) and flies that are null for *dcr2* or *r2d2* have their RNAi activity abolished (Lee et al., 2004b; Liu et al., 2006). For fly miRNAs, it is less clear if there is a formal RLC but fly AGO1 (miRNA-associated AGO) prefers to bind miRNA duplexes that contain central mismatches and these small RNA duplexes are specifically generated from pre-miRNAs by Dicer-1/Loquacious (Forstemann et al., 2005; Forstemann et al., 2007; Tomari et al., 2007). In mammals, there is only one Dicer protein that can generate siRNAs or miRNAs and there appears to be no specificity to having the small RNA duplex interact with one AGO protein over another (mammals have 4 AGO proteins, Figure 1.3

and Appendix A1; Yoda et al., 2010; Bernstein et al., 2001). Unlike RNAi in flies, Dicer processing is uncoupled from the formation of pre-RISC in mammals (Yoda et al., 2010; Betancur and Tomari, 2012). Even dsRNA that is cleaved by mammalian Dicer shows no association with the process of pre-RISC formation and then subsequent mature RISC (Yoda et al., 2010). Both mammals and flies require ATP to form the pre-RISC complex — AGO bound to the RNA duplex (Kobayashi and Tomari, 2015). This process first starts with the selection of one of the strands from the small RNA duplex to become the guide strand associated with the AGO protein. Selection is not a random event; rather it is mediated by the thermodynamic stability of the 5′ ends of small RNA duplex, where the least stable 5′ end is preferentially associated with AGO (Khvorova et al., 2003; Schwarz et al., 2003; Tomari et al., 2004). For fly RNAi, strand asymmetry is determined by the RLC whereas in mammals the strand selection is believed to be mediated by the AGO protein itself and the protein's 5′ phosphate-binding pocket (Tomari et al., 2004; Ma et al., 2005; Parker et al., 2005). For miRNA duplexes in flies and mammals (small RNA duplexes that are typically imperfect and contain central mismatches, Figure 1.1), the strand that has a lower 5′ end thermodynamic stability becomes the guide in AGO (referred to as the miRNA strand) while the strand with a more stable 5′ end does not associate with AGO, the miRNA* strand (Okamura et al., 2009; Yoda et al., 2010; Kobayashi and Tomari, 2015).

The exact mechanism of converting pre-RISC, AGO bound to a small RNA duplex, to become mature RISC, AGO bound to a single-stranded RNA, is not completely clear but the process is one that involves several heat-shock chaperone proteins that hydrolyze ATP (Kobayashi and Tomari, 2015). An early model proposed by Tomari and colleagues attributed the roles of the chaperones to assisting with AGO negotiating the bound RNA duplex in pre-RISC formation, and this model was referred to as the 'rubber band' model (Iwasaki et al., 2010; Kawamata and Tomari, 2010). Recently, ensemble and single-molecule biochemical experiments led Tomari and colleagues to revise this model so that it is in line with new structural data (Iwasaki et al., 2010; Elkayam et al., 2012; Nakanishi et al., 2012; Schirle and MacRae, 2012; Iwasaki et al., 2015a; Iwasaki et al., 2015b). When AGO is loaded with a siRNA or miRNA duplex the action of a passenger strand (complementary to the guide strand in an siRNA duplex) or miRNA* strand removal does not require ATP hydrolysis (Iwasaki et al., 2010; Iwasaki et al., 2015a). The first steps involved with passenger or miRNA* strand removal involve the N-domain of the AGO protein wedging in-between one end of the RNA duplex (Kwak and Tomari, 2012). This wedging activity of the N-domain is thought to disrupt the base pairing past the 16th nucleotide of the guide strand (g16) (guide positions are noted by "g" and the nucleotide number is respective to the 5′ end of RNA; Kwak and Tomari, 2012; Wang et al., 2009). Proper N-domain activity and wedging action is critical

for the subsequent ejection of the passenger or miRNA* strand (Kwak and Tomari, 2012). The final steps in passenger or miRNA* strand removal involve a cleavage-dependent or cleavage-independent process. For miRNAs, they typically contain central mismatches that prevent duplex cleavage by a cleavage-competent AGO proteins (mammalian AGO2 is the only cleavage competent AGO) which, when combined with specific features found in different domains of AGO, lead to the destabilization of the miRNA/miRNA* duplex and eventual disassociation of the miRNA* strand (Hock and Meister, 2008; Kawamata et al., 2009; Yoda et al., 2010; Kawamata et al., 2011). Mammalian AGO1, 3, and 4 along with fly AGO1, do not have in vivo cleaving activity and/or do not contain catalytic residues to be a cleavage-competent AGO protein (Figure 1.3). When a non-cleaving AGO is presented with a siRNA duplex it typically does not form mature RISC or does so very slowly (Kawamata et al., 2009). For fly AGO2 and mammalian AGO2 that have cleavage activity they can cleave the passenger strand of the siRNA duplex in a Mg^{2+}-dependent fashion after wedging (Matranga et al., 2005; Miyoshi et al., 2005; Rand et al., 2005; Shin, 2008). After cleavage of the siRNA duplex, base pairing is now unstable which allows for the cleaved passenger strand to be ejected. It was reported that the mega endonuclease complex, C3PO, made up of an octamer of Trax and Translin proteins degrades the cleaved passenger strand to accelerate mature RISC formation (Liu et al., 2009; Ye et al., 2011).

Figure 1.3

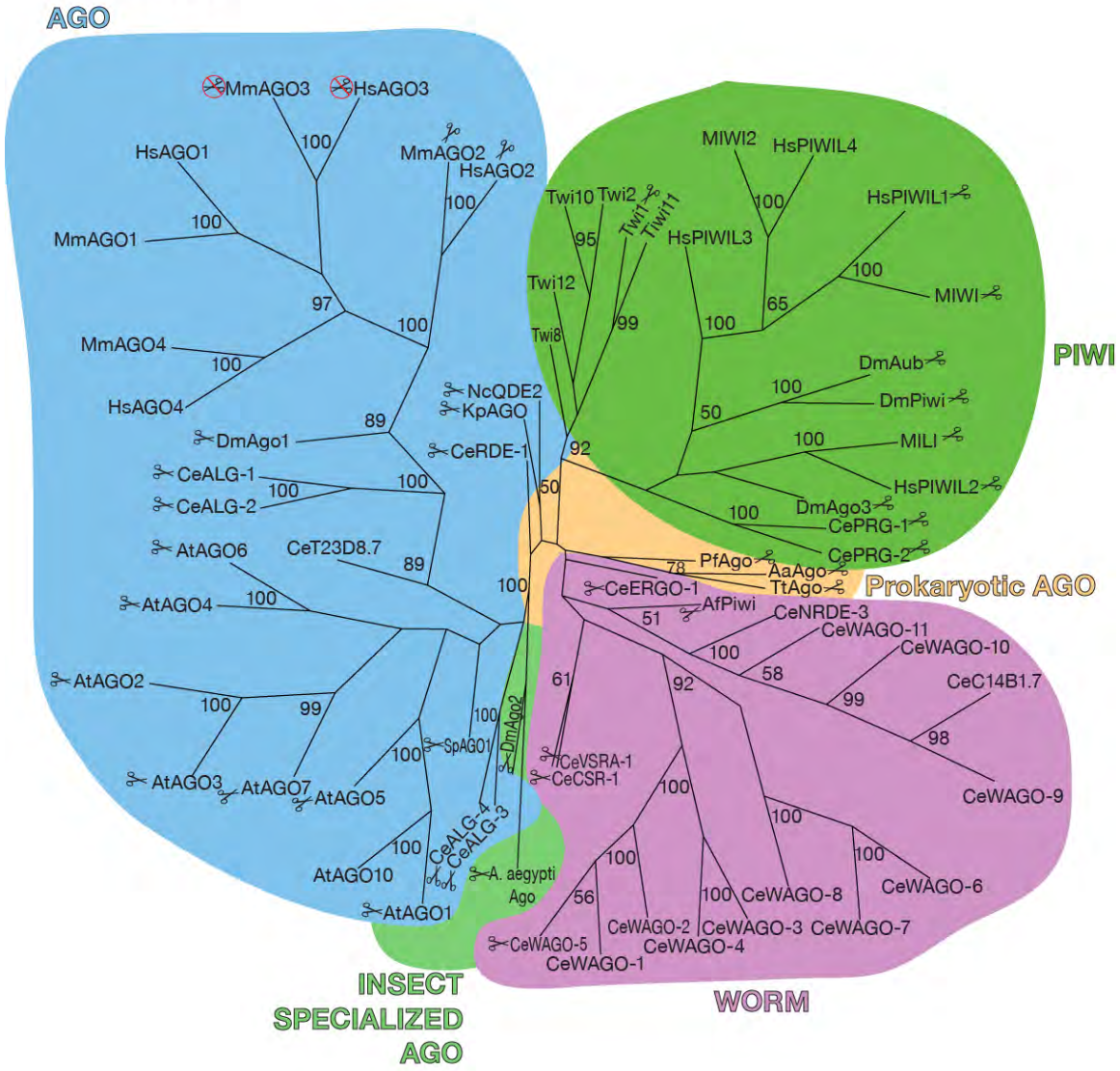

AGO
MmAGO3
HsAGO3
HsAGO1
MmAGO2
HsAGO2
MmAGO1
MmAGO4
HsAGO4
DmAgo1
NcQDE2
KpAGO
CeRDE-1
CeALG-1
CeALG-2
AtAGO6
CeT23D8.7
AtAGO4
AtAGO2
AtAGO3
AtAGO7
AtAGO5
SpAGO1
DmAgo2
CeALG-4
CeALG-3
AtAGO10
AtAGO1
A. aegypti Ago
MIWI2
HsPIWIL4
Twi10
Twi2
Twi1
Twi11
HsPIWIL1
HsPIWIL3
Twi12
MIWI
Twi8
DmAub
DmPiwi
MILI
HsPIWIL2
DmAgo3
CePRG-1
CePRG-2
PIWI
PfAgo
AaAgo
TtAgo
Prokaryotic AGO
CeERGO-1
AfPiwi
CeNRDE-3
CeWAGO-11
CeWAGO-10
CeC14B1.7
CeWAGO-9
CeVSRA-1
CeCSR-1
CeWAGO-6
CeWAGO-8
CeWAGO-2
CeWAGO-3
CeWAGO-7
CeWAGO-4
CeWAGO-5
CeWAGO-1
INSECT SPECIALIZED AGO
WORM

Figure 1.3: Argonaute proteins are categorized into specific clades.

Amino acid sequences of Argonaute proteins were aligned using ClustalX. Phylogenetic tree was generated by taking aligned sequences and using Phylip for bootstrapping, calculation of protein distance, and to generate consensus tree (Larkin et al., 2007; Goujon et al., 2010; Wee, 2013). Bootstrap percentages greater than 50% are indicated at the forks. Aa: *Aquifex aeolicus*, A. aegypti: *Aedes aegypti* Af: *Archaeoglobus fulgidus* At: *Arabidopsis thaliana*, Ce: *Caenorhabditis elegans*, Dm: *Drosophila melanogaster*, Hs: *Homo sapiens*, Kp: *Kluyveromyces polysporus* Mm: *Mus musculus*, Nc: *Neurospora crassa*, Pf: *Pyrococcus furiosus*, Sp: *Schizosaccharomyces pombe*, Tt: *Thermus thermophilus*, Twi: *Tetrahymena* Piwi Argonaute. Scissors represent presence of catalytic activity or catalytic residues within the PIWI domain. Human and mouse AGO3 have the catalytic residues in the PIWI domain but they cannot cleave RNA due to the absence of structural features in the N domain (Meister et al., 2004b; Hauptmann et al., 2013).

It is unclear how this protein complex degrades the short, cleaved passenger strand given its active site is in the center of the protein complex and it is nearly double the size of AGO (Ye et al., 2011). These processes are all ATP-independent however there is a clear requirement for ATP in both miRNA and siRNA RISC maturation. The new model proposed by Tomari and colleagues places the role of the ATP-consuming heat shock chaperone proteins to keeping unloaded AGO proteins in an active state so that they can accept a small RNA duplex rather than assisting in negotiation and/or ejection of the passenger strand (Kobayashi and Tomari, 2015). This model places the chaperones role further upstream than the previous one ('rubber band' model) and equates the ATP hydrolysis by heat shock proteins to making sure AGO remains in an 'active' conformation to accept a duplex that will then lead to the anchoring of the 5′ and 3′ ends of the eventual guide or miRNA strand (Kobayashi and Tomari, 2015). Future structural studies will be needed to confirm this model, in particular a structure of the pre-RISC complex — chaperone proteins associated with AGO as the small RNA duplex binds.

miRNA-mediated target regulation

After the miRNA complexes with an AGO protein to form a microRNA-RISC it is now licensed for post-transcriptional regulation of RNA targets in the cell. The role of miRNA in regulation is integral to many biologic processes and they participate in a network of gene regulation with other miRNA and

RNA-binding proteins (Gurtan and Sharp, 2013; Jonas and Izaurralde, 2015). It is predicted that there are nearly 1,500 miRNAs found in humans and that they can regulate over half of all protein coding genes (Friedman et al., 2009; Chiang et al., 2010). The miRNA-RISC binds to its target through partial complementarity which is typically to nucleotides g2–g8 of the miRNA guide (Lewis et al., 2003; Rajewsky and Socci, 2004; Krek et al., 2005; Lewis et al., 2005; Lim et al., 2005). The AGO protein when bound to a small RNA changes the shape of the small RNA where g2–g8 are arranged in a pre-helical formation to increase target finding efficiency — this is referred to as the 'seed' of the guide strand (described in detail later in this chapter ; Ma et al., 2004; Parker et al., 2005; Wang et al., 2008a; Parker et al., 2009; Elkayam et al., 2012; Schirle and MacRae, 2012; Nakanishi et al., 2012).

When miRNA-RISC binds to its target it acts as a scaffold for other proteins that will regulate the target through destabilization and consequentially repress translation of the mRNA (Figure 1.1; Jonas and Izaurralde, 2015). The GW182 family of proteins (in mammals, Tri-nucleotide repeat-containing proteins, TNRC6A, TNRC6B, and TNRC6C) are one of the co-factors associated with AGO that promotes destabilization of the target (Eulalio et al., 2008; Eulalio et al., 2009; Lian et al., 2009). The GW182 proteins contain an AGO-binding domain (ABD) along with several Tryptophan (W)-containing motifs that bind several protein in the deadenylase complex (Rehwinkel et al., 2005;

Behm-Ansmant et al., 2006; Takimoto et al., 2009; Braun et al., 2011; Chekulaeva et al., 2011; Fabian et al., 2011). GW182 along with AGO becomes a scaffold for destabilizing factors that will shorten the polyA tail which is predicted to lower translation efficiency (Jonas and Izaurralde, 2015). Furthermore, decreasing the polyA tail length reduces the binding of PABP (polyA-binding protein) that interacts with the 5′ end of the mRNA to promote RNA stability (Bernstein et al., 1989; Mangus et al., 2003; Eichhorn et al., 2014). The lost interaction of the 5′ cap and the 3′ end polyA tail accelerates mRNA degradation by decreasing the stability of the cap binding complex and promoting association with the decapping protein complex (Yamashita et al., 2005; Jonas and Izaurralde, 2013). Loss of the cap makes the mRNA susceptible to exonucleolytic enzymes like XRN1 (5′→3′ exonuclease) and thus degradation of the transcript (Braun et al., 2012; Jonas and Izaurralde, 2015).

siRNA-mediated regulation

After a siRNA binds to a catalytically active Argonaute to form a RISC it can then direct endonucleolytic cleavage of its RNA target (Figure 1.1D). Similar to miRNA-RISC, siRNA-RISC uses the seed, g2–g8 of it bound guide to find targets (Haley and Zamore, 2004; Brennecke et al., 2005; Wee et al., 2012). Unlike miRNA-RISC, the guide stand needs to have extensive complementarity to its RNA target in order to cleave its target (Ding et al., 2003; Haley and Zamore, 2004; Martinez and Tuschl, 2004; Schwarz et al., 2006). In mammals,

AGO2 is the only catalytically active AGO protein (Figure 1.3; Meister et al., 2004b; Liu et al., 2004; Hutvagner and Simard, 2008). When AGO2 binds to a target with extensive complementarity it is believed to undergo a protein conformation change that allows for cleavage of the bound RNA target (Wang et al., 2009; Sasaki and Tomari, 2012; Schirle and MacRae, 2012). The idea of a protein conformation preceding target cleavage was described in the 'two-state' model proposed by Tomari and Zamore (Tomari and Zamore, 2005). The 'two-state' model, before the full structure of AGO was solved, was an effort to explain why the 5′ end nucleotides were more important for target binding over nucleotides in the 3′ end (Tomari and Zamore, 2005). Furthermore, it was proposed that the 3′ end of the guide would then release from AGO2's PAZ domain and allow it to base pair with the target as this was shown biochemically lead to more efficient target cleavage (Haley and Zamore, 2004; Tomari and Zamore, 2005; Sasaki and Tomari, 2012; Wee et al., 2012). This model was later supported by the crystal structures of *Thermus thermophilus* Argonaute with and without a target where it was shown that the catalytic residues of AGO that coordinate a magnesium ion to catalyze the cleavage reaction are not properly positioned when AGO is not bound to a target (Wang et al., 2008a; Wang et al., 2009). AGO2-mediated target cleavage is achieved by three different structural features: a catalytic tetrad in the PIWI domain, the PL3 loop also in the PIWI domain, and the N-domain (Boland et al., 2011; Faehnle et al., 2013). The

cleaved target is now subject to degradation by cellular exonucleases and RNA surveillance machinery like XRN-1 (5′→3′ exonuclease) and RRP44 (3′→5′ exonuclease) associated with the exosome complex (Orban and Izaurralde, 2005).

Argonaute — the core of small regulatory RNA pathways

The Argonaute protein is at the core of the RNA silencing machinery, when bound to a small RNA it creates RISC, the complex that is the effector for regulating an RNA target. The superfamily of Argonaute proteins is broken into 3 clades: AGO, PIWI, and worm-specific AGO proteins (Figure 1.3; Hutvagner and Simard, 2008; Cenik and Zamore, 2011). Mammals typically have 8 Argonaute proteins, 4 AGO-like and 4 PIWI-like Argonautes (mice have 3 PIWI-like proteins whereas humans and non-human primates have 4 PIWI-like proteins) (Hock and Meister, 2008). PIWI-like AGO proteins are associated with the piRNA pathway and AGO-like Argonautes are associated with the miRNA and siRNA pathways (Meister, 2013). Mammalian AGO2 can participate in either the miRNA or siRNA pathway since it is catalytically active (Liu et al., 2004; Meister et al., 2004b; Hutvagner and Simard, 2008). Mouse and human AGO2 share about 99% amino acid homology (see Appendix A2 for sequence alignment) where there are only seven amino acids that differ at the N-terminal region of the protein, two of which are closely related. AGO proteins consist of

four distinct domains: N, PAZ, MID, and PIWI (Figure 1.4). Additionally, there are two long linkers (L1 and L2) between the N-PAZ domains and PAZ-MID domains (Figure 1.4; Kuhn and Joshua-Tor, 2013). The structure of Argonaute is well conserved by virtue of the overall structure of each domain between human AGOs and the thermophilic eubacterium *Thermus thermophilus* AGO (Sasaki and Tomari, 2012).

Figure 1.4

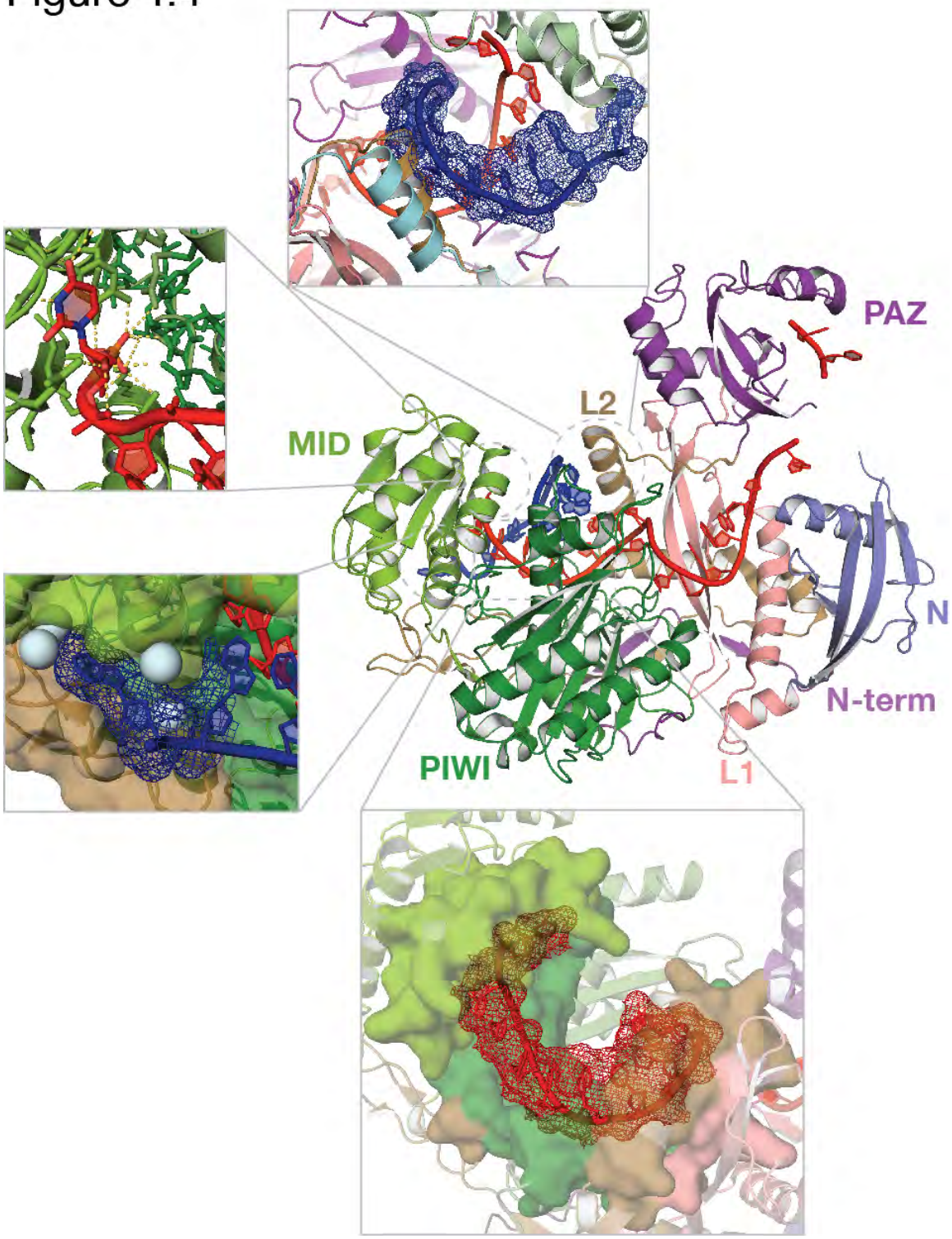


Figure 1.4: Structure of Human AGO2-RISC bound to a short miRNA target

Highlights of some of the prominent structural features of human Argonaute-2. The crystal structure (PDB: 4W5Q, AGO2-RISC with a defined guide strand (red) bound to a short RNA target (blue) that pairs to g2–g8), was solved by Schirle and colleagues (Schirle et al., 2014). Callouts are alignments or displays of the solved structure to highlight detail. Top: α helix 7 of L2 linker moves 4 Å upon binding a miRNA target. Alignment of crystal structures 4W5N (AGO2-RISC bound to no target) and 4W5Q (bound to a target with complementarity to g2–g8) (Schrödinger, 2010). Movement of AGO2-RISC α helix 7 is shown in the unbound state (tan) clashing with the RNA target where as in the target bound state (cyan) α helix 7 moves to accommodate the target. Top left: Nucleotide g1 at the 5′ end of the guide strand is 'flipped out' making contacts within the binding pocket of the MID domain. This nucleotide does not participate in target pairing as it reserved to anchor the 5′ end of the guide strand. Bottom left: Target position 1 (t1) does not pair with the guide strand but adenosine will make water mediated contacts with the MID and L2 domains (Schirle et al., 2015). Bottom center: The pre-helical structure of the seed (target not shown). Argonaute changes the shape of the guide strand to create the seed (g2–g8) and anchors the 5′ end with g1. The half turn of the helix is shown where the bases are pre-organized to pre-pay the entropic penalty associated with nucleic acid hybridization.

N-terminal domain

The N-terminal domain may have been one of the most under appreciated domain of the AGO until several recent structural studies emerged with domain swapping experiments to make non-catalytic AGOs catalytically active (Faehnle et al., 2013; Hauptmann et al., 2013). The first full-length structural studies with *Thermus thermophilus* AGO showed that it blocked duplex propagation between the guide strand and target past the 16th nucleotide, respective to the 5′ end of the guide (Wang et al., 2009). Systematic analysis and biochemical experiments with human AGO2 revealed that the N-domain had an important role in RISC maturation where it acted as a 'wedge' to help pry apart the siRNA duplex for RISC formation (Kwak and Tomari, 2012). The N-domain is required for the cleavage of a siRNA duplex and it is required for duplex unwinding of stable miRNAs duplexes (Kwak and Tomari, 2012). Furthermore, N-domain mutants are not capable of unwinding nicked duplexes from AGO2 or catalytic dead AGO2 (Kwak and Tomari, 2012). This result supports previous loading experiments with AGO1, AGO3, and AGO4 where they all remain in pre-RISC with a siRNA duplexes or nicked duplexes (Yoda et al., 2010; Kobayashi and Tomari, 2015). The N-domain mutants, however, did not have any impact on product release or turnover of AGO2-RISC after target cleavage (Kwak and Tomari, 2012). This observation shows that although duplex removal and product release are similar in the molecular

actions there may be interactions with the target that are distinct from that of interaction with an siRNA or miRNA duplex. The catalytic activity of Argonaute is more than just containing the catalytic tetrad in the PIWI domain, rather it is the combination of the N-domain along with the catalytic tetrad; AGO3 is catalytically inactive for cleaving RNA both in vitro and in vivo even though it contains all of the catalytic residues (Meister et al., 2004b; Hauptmann et al., 2013). Domain swapping revealed that in addition to having a catalytic tetrad in the PIWI domain that mediates Mg^{2+}-dependent cleavage, the N terminal domain plays a critical in slicing activity (Schwarz et al., 2004; Faehnle et al., 2013; Hauptmann et al., 2013). When the N-domain in non-catalytic AGOs is replace with the AGO2 N-domain it then confers catalytic activity in combination with having all catalytic residues along with a conserved Phe residue found in the AGO2 PIWI domain (Faehnle et al., 2013; Hauptmann et al., 2013). The role for the N terminal domain in slicing activity appears to be involved with proper positioning of the guide and target to help align the central cleft of the protein that cleaves target. Further structural studies with AGO2 in the pre-RISC state and when bound to a perfect target will help confirm the insights gleamed from biochemical studies with the N-domain.

PAZ domain

The PAZ (PIWI-Argonaute-Zwille) domain of AGO2 binds RNA and contains 2 subdomains; one with a five-stranded open β-barrel with two helices

on one end of the barrel and an additional strand on the outer part of the barrel, the second subdomain contains a β-hairpin followed by an α-helix (Lingel et al., 2003; Song et al., 2003; Yan et al., 2003). The first subdomain is similar to other single-stranded nucleic acid binding proteins with an oligonucleotide and oligosaccharide-binding fold (Song et al., 2003). The AGO2-PAZ domain when expressed and purified from *Escherichia coli* was reported to have heterogeneous populations RNA, highly suggestive that it is responsible for AGO2's binding to RNA in vivo (Song et al., 2003). Although PAZ binds nucleic acid, in vitro binding assays showed that its K_D is in the low micromolar range whereas full-length AGO2 binds miRNAs at a low nanomolar affinity (Song et al., 2003; Lima et al., 2009; Tan et al., 2009). The PAZ domain of AGO is important for anchoring the 3′ end of the bound guide RNA where it interacts with the 2′ hydroxyl of the terminal nucleotide (Lingel et al., 2004b; Ma et al., 2004; Wang et al., 2008b; Schirle et al., 2014). This differs from the PAZ domain of PIWI proteins that are able to accommodate guides that contain a 3′ end with the terminal nucleotide containing a 2′-*O*-methyl (Simon et al., 2011; Tian et al., 2011).

The PAZ domain in combination with the N-domain creates a narrow channel that the 3′ end of the guide strand threads through from g14–g18 (Schirle et al., 2014). This channel is dynamic as its conformation is impacted when AGO2 binds to a miRNA target (Schirle et al., 2014). After seed pairing, the

channel widens and allows for supplemental pairing to occur between the guide strand and target (if complementarity exists) (Schirle et al., 2014). This widening action of the channel repositions the 3′ half of the guide (g11–g16) to adopt a near A-form conformation that can aid in either supplemental or siRNA-like pairing (Schirle et al., 2014). The A-form conformation is disrupted after g16 or g17 by the N-domain as previously described by the wedging effects of this domain (Wang et al., 2009; Kwak and Tomari, 2012; Schirle et al., 2014). For siRNA targets, the widening of the N-PAZ channel does not just accommodate base pairing but repositions the guide into the central cleft of PIWI to allow for proper coordination of catalytic residues that stimulate endonucleolytic cleavage (Schirle et al., 2014; Kobayashi and Tomari, 2015). It is predicted that the 3′ end of the guide most likely dissociates from the PAZ domain in order to accommodate the target induced protein conformation change (Wang et al., 2009; Schirle et al., 2014; Zander et al., 2014).

PIWI domain

The PIWI domain is the heart of all Argonaute proteins and it is the domain that contains catalytic residues that are required for RNAi along with structural features that change the shape of the bound small RNA (Figure 1.4). A protein domain comparison combined with the structure of prokaryotic Argonaute from *Pyrococcus furiosus* and biochemical identification as AGO2 as the slicer for RNAi led to a redefinition of the AGO protein PIWI domain (Cerutti

et al., 2000; Liu et al., 2004; Song et al., 2004). The core of the PIWI domain contains an RNaseH-like fold with a five-stranded mixed β-sheet that are surrounded by helices and connected to the rest of the protein by an insertion between the last sheet and the helix (Song et al., 2004). The topology of PIWI is similar to other known endonucleases such as: RuvC, a holiday junction endonuclease, Mu and Tn5, transposases, along with the closest match being to the RNase HI and RNase HII enzymes (Katayanagi et al., 1993; Ariyoshi et al., 1994; Rice and Mizuuchi, 1995; Davies et al., 2000; Lai et al., 2000; Song et al., 2004). Like other RNase H enzymes, the PIWI domain coordinates an Mg^{2+} ion in a catalytic center (Yang and Steitz, 1995; Chapados et al., 2001; Liu et al., 2004; Rivas et al., 2005). Mammalian AGO2 uses a catalytic tetrad of Asp-Glu-Asp-His residues to coordinate the Mg^{2+} ion so that it can generate hydroxide ions from water and cause in-line nucleophilic attack on the scissile phosphate of the RNA target across from g10 and g11 of the guide RNA (Martinez and Tuschl, 2004; Schwarz et al., 2004). This cleavage reaction leaves a 5′ phosphate and 3′ hydroxyl on the RNA target, characteristic of other RNase H mediated cleavage reactions (Martinez and Tuschl, 2004; Schwarz et al., 2004).

One of the most important features of the PIWI domain are the interactions it has with the small RNA guide. There are several amino acid contacts (van der Waals, hydrophobic, hydrogen bonds) with the phosphate backbone and sugars of the small RNA guide (Wang et al., 2008b; Elkayam et

al., 2012; Schirle and MacRae, 2012). These interactions act like a ‘cradle’ that the small RNA sits in to form an A-form, pre-helical shape that creates the seed (Figure 1.4; Schirle and MacRae, 2012; Elkayam et al., 2012). One notably difference in the structures of prokaryotic Argonautes and eukaryotic Argonautes is an alpha helical insertion in the PIWI domain that interrupts seed base stacking (Elkayam et al., 2012; Schirle and MacRae, 2012). This structural feature appears to play a role in miRNA target interrogation as it moves approximately 4Å upon target binding (Figure 1.4 ;Schirle et al., 2014). This protein feature is discussed in Chapter III as it relates to the biochemical mechanisms of target finding.

Recently, further comparison of the PIWI domain structure of human AGO1 was solved and compared to human AGO2 (Elkayam et al., 2012; Schirle and MacRae, 2012; Faehnle et al., 2013; Nakanishi et al., 2013). Besides the missing one of the catalytic residues (AGO1 has an Arg instead of a His), there is an insertion segment in the cS7 loop (Faehnle et al., 2013; Nakanishi et al., 2013). The introduction of this loop is predicted to prevent full guide pairing and thus not allowing the RNA target to position correctly for endonucleolytic cleavage (Faehnle et al., 2013; Nakanishi et al., 2013). Creation of AGO1 PIWI domain mutants that create an active catalytic tetrad plus removal of the kink introduced by cS7 loop of AGO1 gave rise to cleavage activity albeit not to the same level of AGO2 (Nakanishi et al., 2013). These results along with AGO2

N domain swapping experiments show an importance of both the PIWI and N domains on cleavage and turnover activity (Faehnle et al., 2013; Hauptmann et al., 2013).

MID Domain

The MID domain contains 4 alternating α-helices and β-strands which resembles a Rossmann fold (Boland et al., 2010; Frank et al., 2010). The β-strands constitute the core of the domain which creates an extended β-sheet that is surrounded by α-helices; the overall architecture of the MID domain is evolutionarily conserved in homologs from human to prokaryotes with some variation in loops and secondary structure (Song et al., 2004; Ma et al., 2005; Wang et al., 2008a; Boland et al., 2010; Frank et al., 2010). The structure of the human MID domain revealed a binding pocket with increased affinity for uracil or adenine over cytosine and guanine (Frank et al., 2010). Different Argonaute proteins show specific nucleotide preferences in the MID domain through specificity loops (Frank et al., 2010). This structural feature helps explain the nucleotide biases in flies and worms where there appears to be specificity for specific small RNAs that begin with one nucleotide over another (Lau et al., 2001; Ghildiyal et al., 2008; Ghildiyal et al., 2010). In addition to nucleotide affinity, the interface between the MID and PIWI domains creates a 5′ phosphate-binding pocket for the small RNA — an anchor point for all small RNAs (Parker et al., 2004). There is a network of amino acid interactions that

secure the 5′ phosphate and in turn the end of the small RNA (Figure 1.4; Wang et al., 2008b; Schirle and MacRae, 2012; Elkayam et al., 2012; Faehnle et al., 2013). These interactions cause the first nucleotide to 'flip-out' and thus it does not participate in base pairing to the target or the pre-helix formed by the seed (Figure 1.4; Wang et al., 2008b; Schirle and MacRae, 2012; Elkayam et al., 2012; Faehnle et al., 2013).

In addition to the interactions the MID domain makes with the small RNA bound to Argonaute, a recent structure of human AGO2 bound to a miRNA target shows specific interaction to the target itself. This data explains computational analyses in vertebrates showing that many miRNA target sites begin with an adenosine at target position 1 even though g1 of the small RNA cannot base pair (Lewis et al., 2005). The structure of human AGO2 with a miRNA target revealed that there is a surface binding pocket between MID and L2 domains of AGO2 that could accommodate the t1 nucleotide of a miRNA target site yet the preference for adenosine was not fully realized (Schirle et al., 2014). The specificity that gives rise to miRNA targets with a t1A is mediated by water and the N6 amine of adenosine that makes specific contacts with the binding cleft on the MID-L2 surface (Figure 1.4; Schirle et al., 2015).

The anatomy of the small RNA guide strand

When a small RNA or DNA binds to an Argonaute it is not just simply a protein-nucleic acid complex. The Argonaute protein changes the shape of the bound

small RNA or DNA to create specific functional domains within the guide strand that contribute to target finding and the overall mechanism of target regulation. The small RNA can be broken into distinct regions: anchor (g1), seed (g2–g8), central (g9–g12), supplementary (g13–g16), 3′ end (g17–g21) (Figure 1.5; Wee et al., 2012).

Figure 1.5

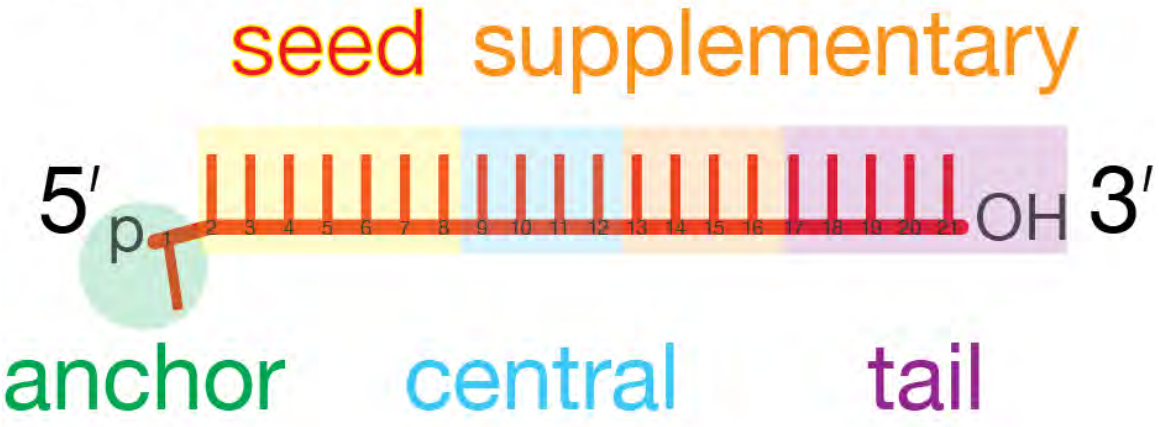

seed
supplementary
5′
p
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
OH
3′
anchor
central
tail

Figure 1.5: Argonaute creates functional domains within its bound small RNA guide.

Anchor

The anchor domain of the small RNA guide interacts with the MID and PIWI domains to secure the 5′ end (Ha and Kim, 2014). This is mediated through a network of amino acid interactions with the phosphate, sugar, and base (Wang et al., 2008b; Elkayam et al., 2012; Schirle and MacRae, 2012; Faehnle et al., 2013). In mammals, there is a preference from adenosine and uracil nucleotides since they have a greater affinity to the MID nucleotide-binding pocket (Bartel, 2009; Frank et al., 2010). Nucleotide preference is mediated by the specificity loops of the MID domain and it allows some organisms (*e.g., D. melanogaster*) to sort their small RNAs into specific Argonautes (Boland et al., 2010; Frank et al., 2010; Ghildiyal et al., 2010; Zha et al., 2012). *Arabidopsis thaliana* Argonautes associate with miRNAs that contains one nucleotide over another but if MID specificity loops are swapped the miRNA will associate with a different Argonaute (Zha et al., 2012). More than just miRNA sorting, the first nucleotide's 5′ phosphate appears to be a most critical anchoring feature (Figures 1.4 and 1.5). Studies with single-stranded RNA guides that bear a 5′ OH instead of a 5′ P were less stable in AGO2 and showed slippage in the cleavage position (Rivas et al., 2005). Kinetic studies show that the anchor (g1) has little to no impact on RISC affinity toward its target and no impact on cleavage activity when it is mismatched to the target (Wee et al., 2012).

Seed

The seed region of the small RNA guide (g2–g8) is responsible for target finding. Structural and kinetic studies have predicted that RISC will find its target faster than nucleic acid alone due to the pre-helical shape the protein puts the RNA in. This pre-helical arrangement and stacking of the bases pre-pays the entropic penalty associated with the hybridization of nucleic acids (Figures 1.4 and 1.5). The seed not only contributes to the affinity for the target but for catalytic Argonautes, seed mismatches could impact the K_M of target cleavage as much as 64-fold (K_M does not equal K_D for Argonautes, $K_D = (k_{cat} + k_{off}) / k_{on}$; Wee et al., 2012).

Central

The central nucleotides of the small RNA guide (g9–g12) are required to base pair for RNAi activity (Ding et al., 2003; Haley and Zamore, 2004; Martinez and Tuschl, 2004; Schwarz et al., 2006). The cleavage of the target takes place at the phosphodiester bond across from g10 and g11 of the guide strand (Elbashir et al., 2001b). Mismatches in this region have the largest impact on k_{cat} in RNAi and a g11/g12 mismatch reduces cleavage activity to non-detectable levels (Wee et al., 2012). Although catalytic activity is greatly impacted by central mismatches, there is little or no change in RISC affinity toward the target — further emphasis that the seed is responsible for target finding and RISC affinity (Ameres et al., 2007; Wee et al., 2012).

Supplementary

The supplementary region pairing (g13–g16) is found in a small subset of miRNAs and is thought to provide increased stability to the miRNA-target interaction, especially with weak seeds (Brennecke et al., 2005; Grimson et al., 2007; Bartel, 2009). The contribution to affinity when compared to the seed region is less important; for example, the let-7a miRNA has no increased affinity over the seed only pairing when compared to seed + 3′ supplementary pairing (Wee et al., 2012). In addition to miRNA modes of regulation, efficient RNAi activity requires base pairing through the central AND supplementary regions (Haley and Zamore, 2004; Wee et al., 2012). Stretches of mismatches through the supplementary region can abolish RNAi activity and similar to seed dinucleotide mismatches, may have pronounced effects on K_M (Wee et al., 2012).

Tail

The tail or 3′ end region (g17–g21) serves as another anchor point for the guide strand interacting with the PAZ domain (Lingel et al., 2003; Song et al., 2003; Yan et al., 2003; Lingel et al., 2004a; Lingel et al., 2004b). There are specific interactions with the 3′ terminal nucleotide of the guide strand. In mammals, there is no terminal 2′-*O*-methyl modification of the sugar, which allows for hydrogen bond interactions to take place at both the 2′ and 3′ positions (Elbashir et al., 2001a; Hutvágner et al., 2001; Wang et al., 2008b;

Elkayam et al., 2012; Faehnle et al., 2013). Additionally, there are van der Waals interactions with the sugar that help it stack along the amino acids (Elkayam et al., 2012; Faehnle et al., 2013; Schirle et al., 2014). These interactions are thought to protect the small RNA from exonucleases as nuclease treatment of AGO2 bound to a small RNA protects it from micrococcal nuclease digestion (De et al., 2013). Similar to other non-seed regions of the guide strand, the tail region has little to no impact on RISC affinity toward its target and mismatches show no changes in K_D (Haley and Zamore, 2004; Wee et al., 2012). Mismatches between the tail and target have been reported to increase product release and subsequently k_{cat} after RNAi, although in some cases this increase is modest (Haley and Zamore, 2004; Wee et al., 2012; De et al., 2013). The sequence composition of the 3′ end tail may play a larger role on the product release step, especially with the N domain acting as a wedge to prevent base pairing past g18 (Kwak and Tomari, 2012; Faehnle et al., 2013; Hauptmann et al., 2013).

Nucleic Acid Hybridization

An underlying biophysical interaction in the mechanism of small regulatory RNAs is the association and dissociation between RNA targets. RISC uses nucleic acid hybridization to discriminate target sites in the transcriptome in order to regulate a complementary target. The principles of both nucleic acid-nucleic acid and protein-nucleic acid interactions factor into the mechanistic workings of small regulatory RNAs.

Numerous studies have looked into the hybridization rates of nucleic acids in a variety of solvents, temperatures, and ionic strengths (Wetmur, 1976). Under physiologic conditions, the temperature and ionic strength strongly favor for nucleic acids to hybridize (Herschlag, 1991). The rate of RNA and/or DNA hybridization is limited by the rate of successful collisions that convert into stable binding events (Ross and Sturtevant, 1960; Ross and Sturtevant, 1962; Nygaard and Hall, 1964; Wetmur and Davidson, 1968). Fundamental studies of nucleic acid association and dissociation have relied upon the spectral properties of nucleic acid where the absorbance at 260 nm increases when denatured and decreases when the bases are stacked in a helix (renatured), this is known as hyperchromicity and hypochromicity, respectively (Wetmur, 1976). The study of renaturation kinetics revealed that this reaction is second-order and concentration dependent (Hutton and Wetmur, 1973; Wetmur, 1976). An exception is when one strand is in way excess of the other strand, this makes the reaction pseudo first-order (Galau et al., 1977). The reaction is not diffusion-controlled, where there are many non-productive (non-helix forming) collisions that precede a productive nucleation event that will lead to the fast ‘zippering’ of the nucleic acid to form the helix (Figure 1.6; Chang et al., 1974; Wetmur and Davidson, 1968; Berg and von Hippel, 1985). The reaction of nucleic acid hybridization is slow, where in addition to non-productive inter-molecular associations there are non-productive intra-molecular events.

The entropy of nucleic acid bases (as illustrated in Figure 1.6) causes a thermodynamic penalty for productive association whereas it is inferred that seed region created by AGO and its bound small RNA or DNA pre-pays this penalty (Wee et al., 2012).

Figure 1.6

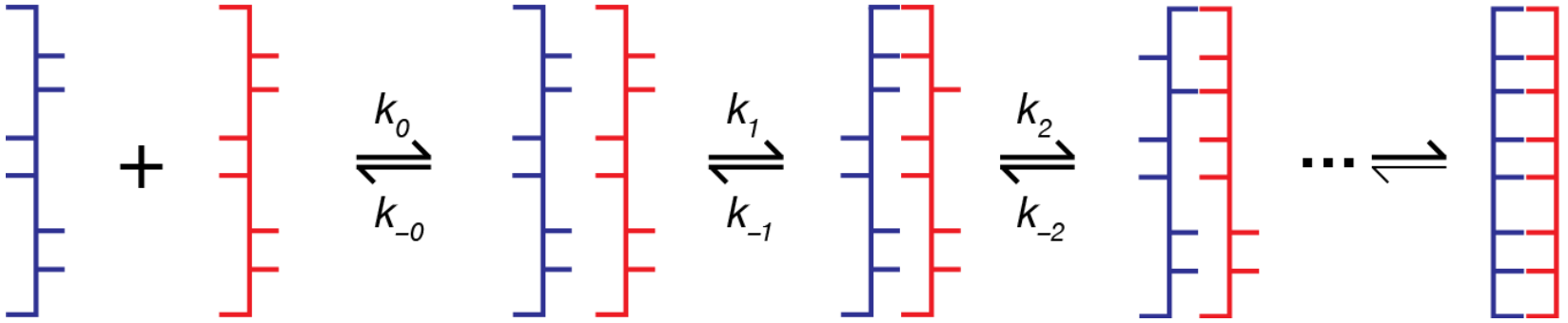


Figure 1.6: Representation of Nucleic Acid Hybridization

Prior to duplex formation, there are many non-productive binding events between single strands of RNA or DNA. The entropy of the individual nucleotides combined intra-molecular interactions contributes to the characteristic slow association kinetics for nucleic acid. When several nucleotides are oriented correctly for base pairing this decreases the dissociation rate of the two strands so that they may ‘zipper’ and form a duplex. The reverse reaction after duplex formation is very slow under physiologic conditions and temperatures.

Once a duplex is formed there are well-established thermodynamic parameters that can be used to predict the stability at a given temperature and/or ionic strength (Xia et al., 1998). The Nearest-neighbor model is based upon empirical measurements of the stability of single base pairs within a duplex relative to the neighboring nucleotides and/or position within the duplex (i.e., terminal nucleotides, bulges, adjacent to an overhang, etc., Nelson et al., 1981; Sugimoto et al., 1987; Longfellow et al., 1990; O'Toole et al., 2005; Xia et al., 1998). These predictions take into account a nucleic acid-nucleic acid interaction but do not and were not intended to take into account how a protein can influence nucleic acid pairing. In the case of small regulatory RNAs, Wee and colleagues showed that different AGO proteins have distinct kinetic properties that can influence RNAi activity and binding kinetics and some of these interactions cannot be predicted by the Nearest-neighbor model alone (Wee et al., 2012; Wee, 2013).

Objective

The underlying mechanism of small regulatory RNAs begins and ends with two simple reaction steps: (1) association/finding its target (2) dissociation after finding its target — this is either releasing cleaved products (RNAi) or dissociating after acting as a scaffold for regulatory silencing factors (miRNA). Some of these steps have been measured in ensemble in vitro kinetic studies or inferred by structural predictions. The goal of this thesis is to understand the

initial and terminal reaction steps of AGO-RISC, the effector complex in small RNA regulation, by direct and quantitative measurement. These measurements utilize a powerful technique that takes measurements of individual molecules in a population rather than the measured average of many molecules in a population. Direct measurement using single-molecule TIRF microscopy coupled to co-localization analysis allows us to measure the interaction of RNA and protein simultaneously. This work highlights the remarkable ability of proteins to discriminate true substrates over mimics and illustrates how efficient enzymes favor substrates over products. This thesis uncovers how the biophysical properties of one macromolecule (small regulatory RNA) can be influenced or changed when bound to another macromolecule (protein, Argonaute). My work shows how components of the small RNA pathway have evolved to suit their regulatory functions in the cell and also highlight a potential insight into how other nucleic acid guided endonucleases parallel Argonaute proteins.

CHAPTER II: Rapid and specific purification of Argonaute-small RNA complexes from crude cell lysates

Disclaimer

This work is the joint effort among the authors: Carlos Fabián Flores-Jasso (CFF-J), William E. Salomon (WES), and Phillip D. Zamore (PDZ). CFF-J performed experiments related to fly RISC and the initial optimization of the purification technique. WES established a mammalian S100 extract that was robust at recapitulating RNAi and optimized the conditions for mouse RISC purification. WES performed experiments, analyses, and prepared figures related to mouse AGO2-RISC. CFF-J and PDZ wrote the manuscript. All authors provided critical review of the data and manuscript. This chapter includes several pieces of unpublished data and some discussion related to the mammalian S100 extract optimization and the purification of mouse AGO2-RISC.

SUMMARY

Small interfering RNAs (siRNAs) direct Argonaute proteins, the core components of the RNA-induced silencing complex (RISC), to cleave complementary target RNAs. Here, we describe a method to purify active RISC containing a single, unique small RNA guide sequence. We begin by capturing RISC using a complementary 2′-*O*-methyl oligonucleotide tethered to beads. Unlike other methods that capture RISC but do not allow its recovery, our strategy purifies active, soluble RISC in good yield. The method takes advantage of the finding that RISC partially paired to a target through its siRNA guide dissociates >300 times faster than a fully paired siRNA in RISC. We use this strategy to purify fly Ago1- and Ago2-RISC, as well as mouse AGO2-RISC. The method can discriminate among RISCs programmed with different guide strands, making it possible to deplete and recover specific RISC populations. Endogenous microRNA:Argonaute complexes can also be purified from cell lysates. Our method scales readily and takes less than a day to complete.

INTRODUCTION

Small interfering RNAs (siRNAs) and microRNAs (miRNAs) mediate the post-transcriptional repression of complementary target RNAs (Bartel, 2009; Ghildiyal and Zamore, 2009). siRNAs typically guide the cleavage of extensively complementary RNAs (Hammond et al., 2000; Zamore et al., 2000; Elbashir et al., 2001a; Elbashir et al., 2001b), a phenomenon called RNA interference (RNAi). In contrast, most animal miRNAs target partially complementary RNAs, triggering their destruction (Baek et al., 2008; Selbach et al., 2008; Ingolia et al., 2009; Guo et al., 2010) and, in some cases, repressing their translation into protein (Hendrickson et al., 2009; Bazzini et al., 2012). Both siRNAs and miRNAs assemble with Argonaute proteins to form the RNA-induced silencing complex (RISC). Assembly of siRNA or miRNA/miRNA* duplexes into RISC is facilitated by proteins that orient the siRNA, such as the Dicer-2/R2D2 heterodimer in insects, and by proteins thought to allow conformational rearrangement of Argonaute, such as HSP90 and Hsc70 chaperones (Tahbaz et al., 2001; Iki et al., 2010; Iwasaki et al., 2010; Johnston et al., 2010; Miyoshi et al., 2010; Olivieri et al., 2012; Preall et al., 2012; Xiol et al., 2012).

In *Drosophila melanogaster,* miRNAs and siRNAs are sorted between Argonaute1 (Ago1) and Argonaute2 (Ago2) according to their duplex structure and the identity of their first nucleotide (Forstemann et al., 2007; Tomari et al., 2007; Czech et al., 2008; Ghildiyal et al., 2008; Okamura et al., 2008b; Seitz et

al., 2011). The extensively double-stranded structure of siRNA sends them into Ago2, while the presence of specific mismatches within the miRNA/miRNA* duplex directs miRNAs into Ago1. In mammals, whose four Argonaute proteins are more closely related to fly Ago1 than to fly Ago2 (Tolia and Joshua-Tor, 2007), no sorting mechanism has thus far been detected, and miRNAs and siRNAs assemble into AGO1, AGO2, AGO3, and AGO4. In mice and humans, only AGO2 can catalyze target cleavage (Liu et al., 2004; Meister et al., 2004b).

Chemically modified, complementary oligonucleotides that bind siRNAs and miRNAs block their activity in vitro and in vivo (Hutvagner et al., 2004; Meister et al., 2004a; Krutzfeldt et al., 2005; Leaman et al., 2005; Elmen et al., 2008; Lanford et al., 2010; Obad et al., 2011). The most widely used anti-miRNA oligonucleotides employ 2′-*O*-methyl ribose modifications to block degradation of the oligonucleotide. Such 2′-*O*-methyl anti-miRNA oligonucleotides not only inhibit miRNA function in cultured cell lines, cell lysates, and *Caenorhabditis elegans* (Hutvagner et al., 2004; Zheng et al., 2010), but also can be used to capture Argonaute complexes programmed with a specific small RNA sequence (Hutvagner et al., 2004). Unfortunately, RISC captured using fully complementary 2′-*O*-methyl oligonucleotides tethered to beads cannot be recovered under native conditions, precluding its further study.

Despite great advances in understanding the miRNA and siRNA pathways, the biochemical details of how Argonaute proteins function remains

incomplete. A key limitation in the study of Argonaute protein function is the lack of methods to purify RISC complexes assembled through natural pathways and that contain a single, unique guide sequence. Purification of RISC using antibodies against endogenous or epitope-tagged Argonautes allows the selection of specific Argonaute proteins, but these contain a complex mixture of siRNA and miRNA guides (Hutvagner and Zamore, 2002; Mourelatos et al., 2002; Liu et al., 2004; Meister et al., 2004a; Ikeda et al., 2006; Beitzinger et al., 2007; Azuma-Mukai et al., 2008). RISC has also been purified using a guide strand with a 3′ biotin joined to the siRNA through a UV-sensitive linker, which was cleaved by photolysis (Martinez et al., 2002; Martinez and Tuschl, 2004; Ameres et al., 2007) and using tethered siRNAs from which proteins were recovered with denaturing buffers (Gerbasi et al., 2010). However, photolysis is inefficient, and recovery by denaturation, of course, fails to preserve RISC activity. Both approaches can only be used to isolate experimentally programmed RISC, whereas our method permits purification of endogenous complexes from cell and tissue extracts.

Recombinant eukaryotic Argonaute proteins have been produced in bacteria, and in insect, yeast, human, and lepidopteran cells (Rivas et al., 2005; MacRae et al., 2008; Ye et al., 2011; Elkayam et al., 2012; Nakanishi et al., 2012; Schirle and MacRae, 2012). Recombinant Argonaute proteins can be partially inactive (Rivas et al., 2005) and often contain endogenous RNAs (Nakanishi et

al., 2012; Schirle and MacRae, 2012). Furthermore, recombinant Argonaute proteins must be loaded with single-stranded RNA, a pathway not thought to exist in vivo, and lack associated proteins that may modify Argonaute function.

Here, we describe a method to purify active Argonaute complexes (RISCs) containing a single, unique small RNA guide. Like previous techniques, our method begins by capturing RISC using a complementary 2′-*O*-methyl oligonucleotide tethered to paramagnetic beads (the capture oligo). The strategy takes advantage of the finding that fly Ago2-RISC dissociates from a target RNA >300 times more rapidly when its guide strand is partially paired than when it is fully paired to a target (Wee et al., 2012). We can therefore elute purified, active, soluble RISC from the capture oligo in good yield in native conditions. We use this strategy to purify *Drosophila* Ago2-RISC bearing an siRNA and Ago1-RISC loaded with a miRNA, as well as an siRNA assembled into mouse AGO2 in lysates from immortalized embryonic fibroblasts. Finally, we show that active RISC can be selectively purified from a population of RISCs containing different guide strands. This method scales readily and takes no more than a single day to complete.

RESULTS

miRNA-like pairing allows efficient capture and release of RISC

We incubated an siRNA duplex in *Drosophila* embryo lysate to assemble the guide strand into Ago2-RISC. We captured the RISC containing the siRNA guide using a 31 nt 2′-*O*-methyl oligonucleotide tethered via a 5′ biotin to streptavidin-paramagnetic beads. RISC, captured on the oligo, was then washed in buffer containing 2 M potassium acetate and eluted with a competitor DNA oligonucleotide fully complementary to the capture oligo in buffer containing 1 M potassium acetate. Figure 2.1 illustrates the method. Three types of capture oligos were tested: an oligo fully complementary to the siRNA guide; an oligo complementary only to the seed of the siRNA (positions 2–8); and an oligo complementary to the seed and four nucleotides in the 3′ supplementary region of the siRNA. The amount of capture oligo used (5 μM) was greater than the concentration of 2′-*O*-methyl oligonucleotide (1–2 μM) that fully blocks target cleavage directed by this siRNA in *Drosophila* embryo lysate (Figure S2.1).

Figure 2.1

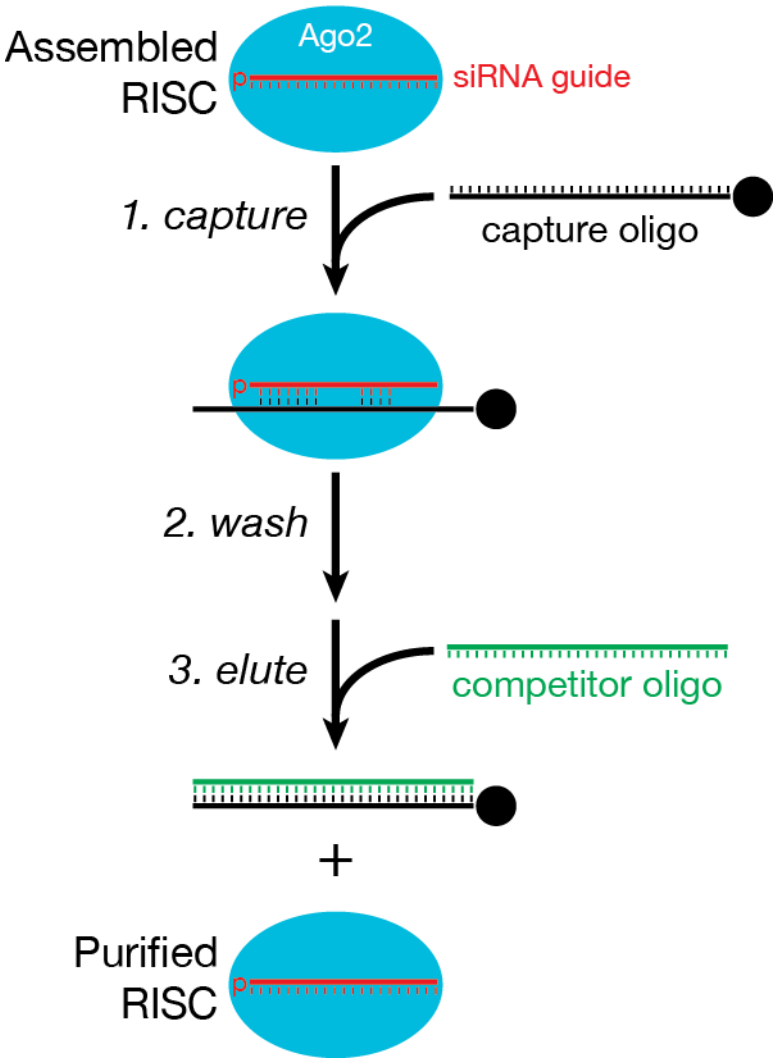

Assembled
RISC
Ago2
siRNA guide
1. capture
capture oligo
2. wash
3. elute
competitor oligo
+
Purified
RISC

Figure 2.1: Strategy to purify active RISC

To monitor the efficiency of binding and recovery of siRNA-programmed RISC for each of the three types of capture oligos, we assayed cleavage of a 5′ ^{32}P-radiolabeled target RNA (Figure 2.2). In parallel, we followed a 5′ ^{32}P-radiolabeled guide siRNA through the purification procedure (Table 2.1). The fully complementary capture oligo, the capture oligo with seed plus 3′ supplementary pairing, and the capture oligo pairing only with the seed, all effectively depleted the programmed RISC from the lysate: we detected less than 2% of the original RISC cleaving activity in the supernatant of an assembly reaction incubated with any of the capture oligos tethered to paramagnetic beads or in the subsequent, pooled washes (100 mM and 2 M potassium acetate). For the fully complementary capture oligo, addition of a competitor oligonucleotide (the competitor) complementary to the capture oligo failed to release any detectable target-cleaving RISC activity (Table 2.1). In contrast, a competitor efficiently released active RISC from the capture oligo complementary to only the seed plus the 3′ supplementary region or the seed alone. We note that high salt (1 M potassium acetate) was essential: at lower salt concentrations, addition of a competitor complementary to the capture oligo failed to release any detectable target-cleaving RISC activity (Figure 2.2) or ^{32}P-radiolabeled guide siRNA (Table 2.1, compare “low salt + competitor” to “eluate”).

Figure 2.2

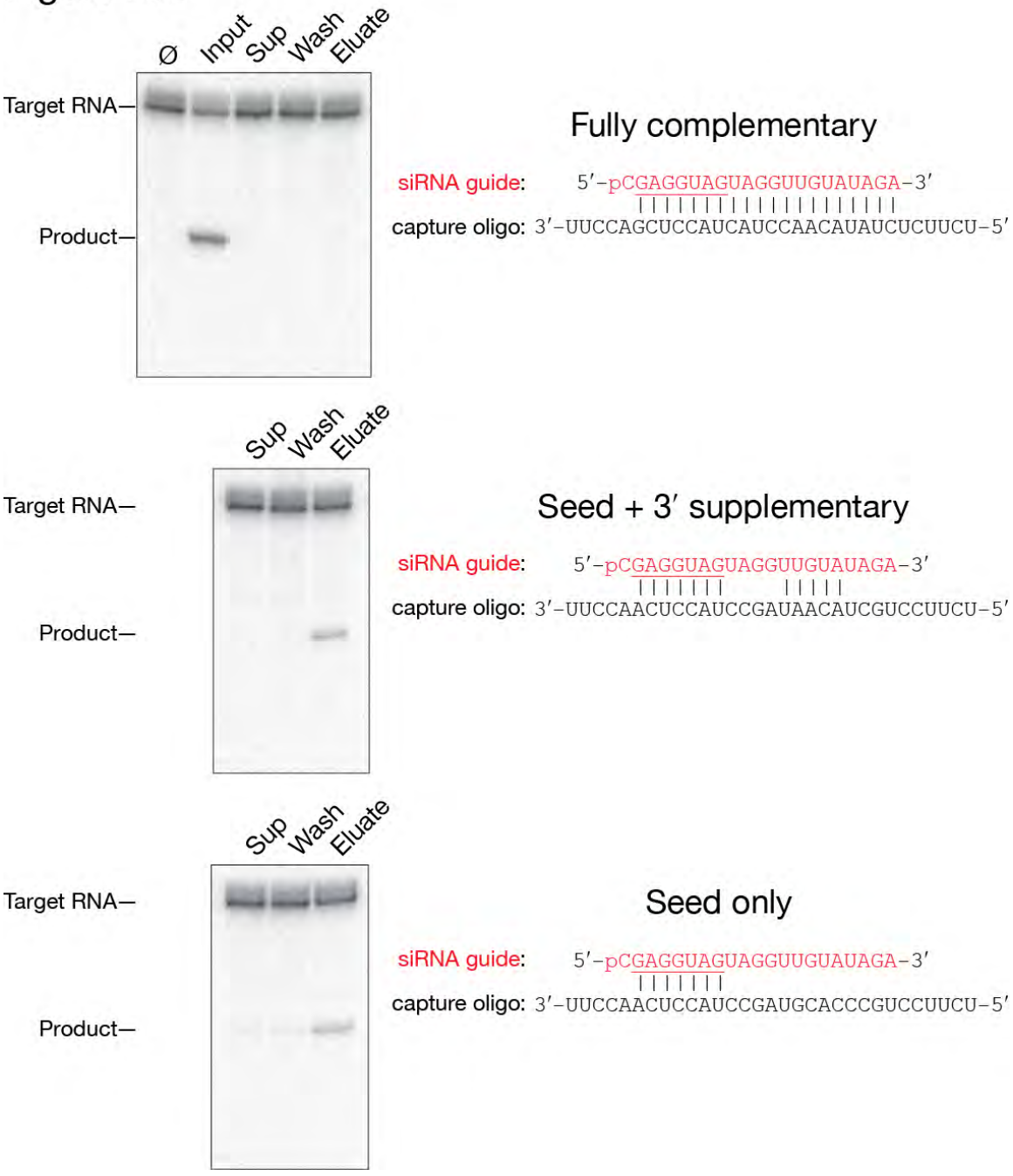

Ø
Input
Sup
Wash
Eluate
Target RNA—
Product—
Fully complementary
siRNA guide: 5′-pCGAGGUAGUAGGUUGUAUAGA-3′
capture oligo: 3′-UUCCAGCUCCAUCAUCCAACAUAUCUCUUCU-5′
Sup
Wash
Eluate
Target RNA—
Product—
Seed + 3′ supplementary
siRNA guide: 5′-pCGAGGUAGUAGGUUGUAUAGA-3′
capture oligo: 3′-UUCCAACUCCAUCCGAUAACAUCGUCCUUCU-5′
Sup
Wash
Eluate
Target RNA—
Product—
Seed only
siRNA guide: 5′-pCGAGGUAGUAGGUUGUAUAGA-3′
capture oligo: 3′-UUCCAACUCCAUCCGAUGCACCCGUCCUUCU-5′

Figure 2.2: Active RISC can be eluted from a capture oligo partially complementary to the small RNA guide

let-7-programmed *Drosophila* Ago2-RISC, assembled in in *Drosophila* embryo lysate, was incubated with a capture oligo fully complementary to *let-7*, to the *let-7* seed plus 3′ supplementary region, or only to the *let-7* seed sequence. The RISC assembly reaction (Input), the supernatant after capture (Sup), the washes (five with 100 mM and five with 2 M potassium acetate, pooled and concentrated to the volume of the original RISC assembly reaction, 100 µl; Wash), and the eluate from the capture oligos were incubated with 100 nM target RNA for 5 min at 25°C to detect Ago2-RISC activity. Ø, no incubation.

TABLE 2.1

Capture oligo complementarity	Active RISC assembled	siRNA in supernatant	Washes			Eluate (high salt + competitor)	Yield
			Low salt	Low salt + competitor	High salt		
None	12 ± 1	16 ± 1	0.15 ± 0.01	0.10 ± 0.01	0.83 ± 0.06	0.13 + 0.01	1.1%
Complete complementarity to guide strand	12 ± 1	10 ± 1	0.43 ± 0.01	0.092 ± 0.008	0.10 ± 0.01	0.29 ± 0.03	2.4%
Guide strand seed plus 3′ supplementary	12 ± 1	7.4 ± 0.4	0.10 ± 0.02	0.12 ± 0.02	0.10 ± 0.02	9.4 ± 0.1	78%
Guide strand seed only	12 ± 1	8.1 ± 0.4	0.11 ± 0.01	0.11 ± 0.01	0.097 ± 0.009	9.6 ± 0.6	80%
Passenger strand seed plus 3′ supplementary	12 ± 1	21 ± 1	0.22 ± 0.05	0.12 ± 0.07	0.082 ± 0.008	0.082 ± 0.008	0.68%

Table 2.1: Yield of fly Ago2-RISC eluted from the capture oligo

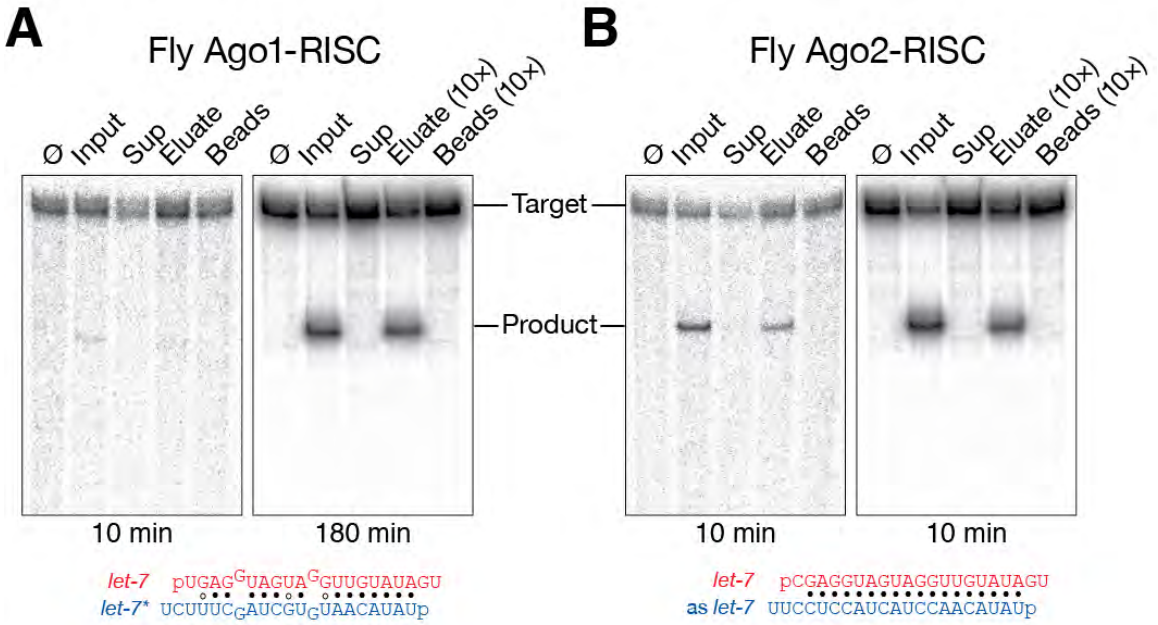
RISC was assembled in Drosophila embryo lysate using 25 nM siRNA duplex. For the supernatant and washes, the amount of 5′ ^{32}P-radiolabeled siRNA (pmol) was measured as a surrogate for RISC. Active RISC assembled and eluted were measured as described in the Materials and Methods. Data are mean ± S.D., n = 3.The RISC eluted by this method contains competitor DNA that is partially complementary to an RNA target fully matching the siRNA guide. In practice, the contaminating competitor DNA does not detectably interfere with target cleavage. However, the competitor can be readily removed by subsequent ion exchange chromatography (See Materials and Methods).

Purification of Ago1 programmed with a miRNA

In flies, canonical miRNA/miRNA* duplexes load Ago1 rather than Ago2 (Forstemann et al., 2007; Tomari et al., 2007). To test whether the method could be used to selectively purify Ago1-RISC, we incubated a *let-7*/*let-7** duplex with *Drosophila* embryo lysate, and then purified the *let-7*-programmed Ago1-RISC using the capture oligo complementary to the seed plus the 3′ supplementary region. Although Ago 1-RISC is inherently less catalytically active than Ago2 ($kcat/K_M$ for Ago2 is >60-fold greater than for Ago1) and target cleavage is unlikely to play an important role in Ago1-mediated mRNA repression (Forstemann et al., 2007), Ago1 can be followed by target cleaving activity (Figure 2.3A). Target cleavage assays of *let-7*-programmed Ago1-RISC detected activity in the input and eluate samples (Figure 2.3A). Quantitative mass spectrometry (Figure 2.3C) revealed that the eluate was more enriched for Ago1 (467-fold over background) than for Ago2 (fivefold enrichment). In contrast, RISC assembled with a *let-7* siRNA (Figures 2.3B, 2.3C and Tables S2.1 and S2.2) contained mainly Ago2 (356-fold enrichment over background) rather than Ago1 (1.6-fold enrichment). Importantly, we used the same capture oligo to purify each Argonaute complex, changing only the *let-7* duplex used to assemble RISC. Flies do not produce *let-7* during embryogenesis (Pasquinelli et al., 2000),

so all *let-7* RISC activity reflects Argonaute programmed in the lysate by the *let-7* siRNA or miRNA/miRNA* duplex.

Figure 2.3


A
Fly Ago1-RISC
Ø Input Sup Eluate Beads
Ø Input Sup Eluate (10x) Beads (10x)
Target
Product
10 min
180 min
let-7 pUGAGGUAGUAGGUUGUAUAGU
let-7* UCUUUCGAUCGUGUAACAUAUp
B
Fly Ago2-RISC
Ø Input Sup Eluate Beads
Ø Input Sup Eluate (10x) Beads (10x)
10 min
10 min
let-7 pCGAGGUAGUAGGUUGUAUAGU
as let-7 UUCCUCCAUCAUCCAACAUAUp

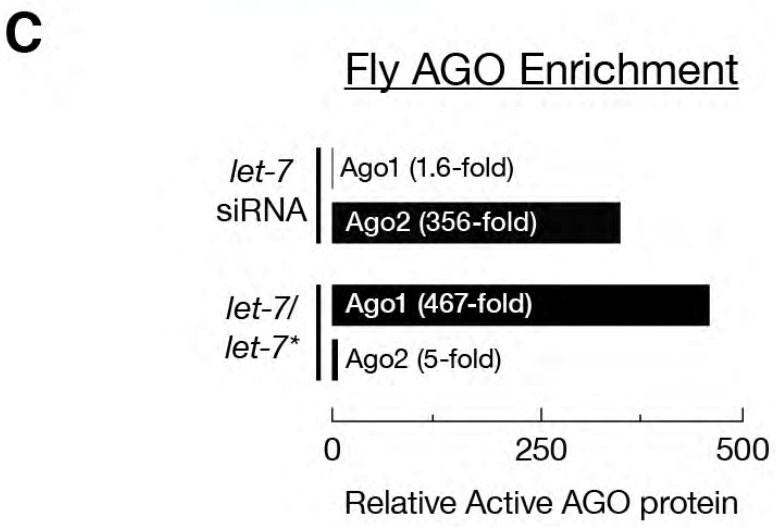

C
Fly AGO Enrichment
let-7 siRNA
Ago1 (1.6-fold)
Ago2 (356-fold)
let-7/ let-7*
Ago1 (467-fold)
Ago2 (5-fold)
0
250
500
Relative Active AGO protein

Figure 2.3: *Drosophila* Ago1- and Ago2-RISC, as well as mouse AGO2-RISC, can be purified using a partially complementary capture oligo.

(A) A *let-7*/*let-7** duplex was loaded into *Drosophila* Ago1 or (B) a *let-7* siRNA was loaded into *Drosophila* Ago2 by incubation in *Drosophila* embryo lysate and purified with a partially complementary capture oligo. The RISC assembly reaction (Input), the supernatant after capture (Sup), the first wash, and the eluates were incubated with 100 nM *let-7* complementary target RNA for 10 min (Ago1 and Ago2) and 180 min (Ago1) at 25°C. (C) Tenfold concentrated, purified *Drosophila* Ago1- and Ago2-RISC and control samples were analyzed by quantitative mass spectrometry to determine their protein composition. The enrichment of each Argonaute protein in the purified RISC was calculated as the ratio of *let-7*-programmed, purified samples to control samples in which the small RNA duplex was omitted (background).

Purification of mouse AGO2-RISC

We developed a cell extract system that robustly recapitulates RNAi in vitro that is superior to commonly used HeLa cell extracts for mammalian RNAi (Figures 2.4 and S2.2). The improved cell extract is cytosolic S100 extract from immortalized, *Ago2*$^{-/-}$ mouse embryonic fibroblasts (MEFs; Broderick et al., 2011; O'Carroll et al., 2007) that stably express about 30-fold more *Ago2* mRNA than wild-type MEFs and contain ~200-fold more AGO2 in the S100 cytosolic extract (Figure 2.4). Our method also effectively purifies mammalian AGO2-RISC. The S100 was incubated with a *let-7* siRNA duplex to program RISC and we used the same capture oligo and protocol as for fly RISC, and then tested the eluate for target cleaving activity (Figure 2.4D). Again, target cleaving activity was detected in the eluate when we used a capture oligo complementary to seed plus the 3′ supplementary region or to the seed alone, but not the fully complementary capture oligo.

Figure 2.4

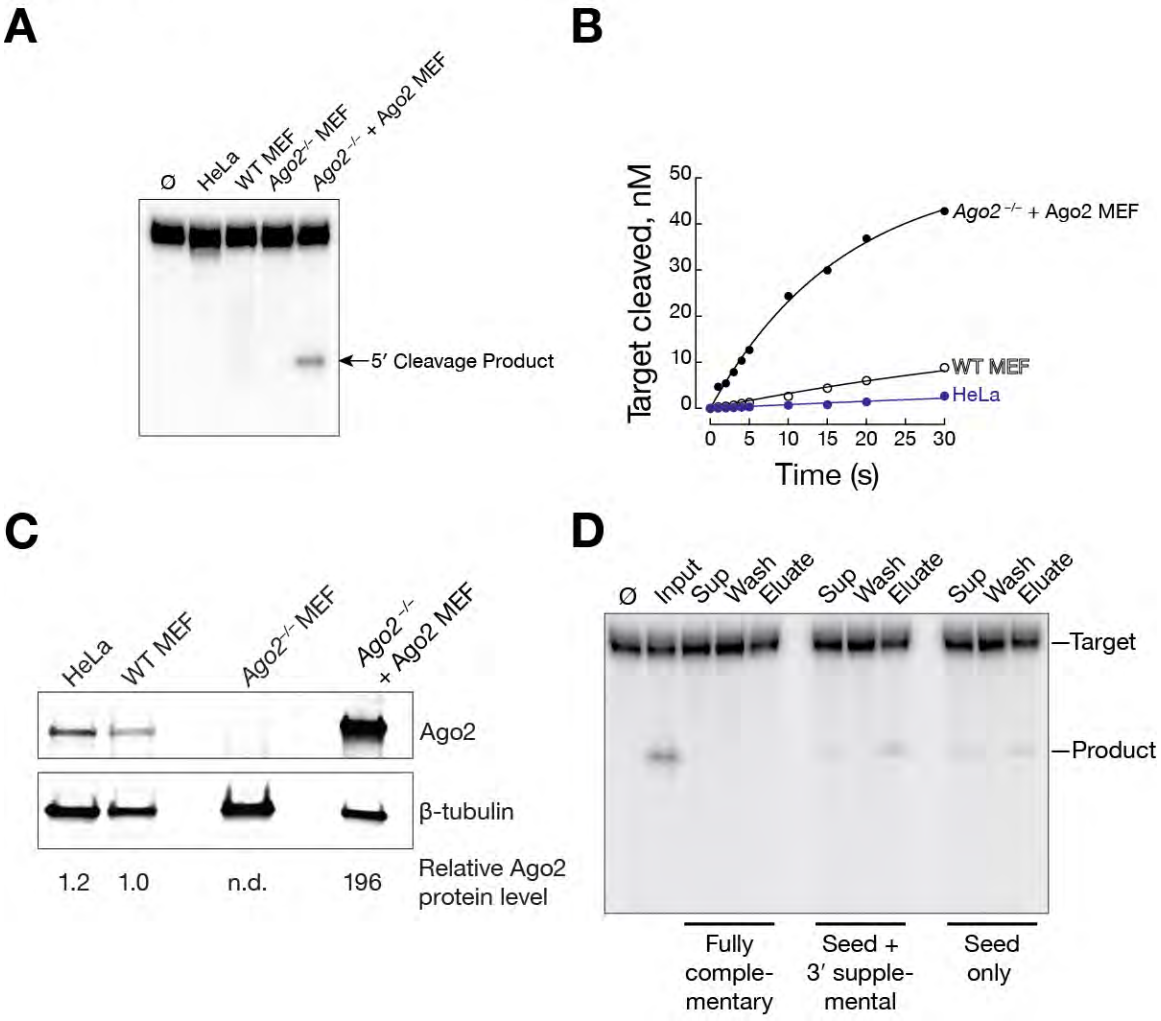

A
Ø
HeLa
WT MEF
Ago2−/− MEF
Ago2−/− + Ago2 MEF
5′ Cleavage Product
B
Target cleaved, nM
50
40
30
20
10
0
0 5 10 15 20 25 30
Time (s)
Ago2 −/− + Ago2 MEF
WT MEF
HeLa
C
HeLa
WT MEF
Ago2−/− MEF
Ago2 −/− + Ago2 MEF
Ago2
β-tubulin
1.2 1.0 n.d. 196
Relative Ago2 protein level
D
Ø
Input
Sup
Wash
Eluate
Sup
Wash
Eluate
Sup
Wash
Eluate
Target
Product
Fully comple-mentary
Seed + 3′ supple-mental
Seed only

Figure 2.4: Establishment of a mammalian in vitro RNAi extract system and purification of mouse Ago2

(A) Representative target cleavage for different S100 cytosolic cell extracts programmed with *let-7* siRNA after 5 minutes incubation with *let-7* target. S100 extracts and target cleavage assay were performed as described in Material and Methods. (B) Quantification of RNAi target cleavage guided by *let-7* siRNA programmed in mammalian S100 cell extracts. (C) Western blot of AGO2 levels in S100 cytosolic extracts. Stoichiometric amounts of total protein (37.5ug) were loaded on a 4–20% SDS-polyacrylamide gel with the exception of S100 extract from *Ago2*$^{-/-}$ MEFs over-expressing mouse AGO2 (20ug) in order to prevent overexposure of the membrane. The relative amount of Ago2 was calculated by normalizing loading to β-tubulin and comparing to wild-type MEF S100 (set at 1.0). (D) Purification of *let-7*-programmed mouse AGO2-RISC. *let-7* siRNA was assembled in S100 cytosolic extract from *Ago2*$^{-/-}$ MEFs over-expressing mouse AGO2 then purified using either fully or partially complementary capture oligos. Target cleaving activity was tested using 100 nM target RNA for 5 min at 37°C.

Although MEFs express at least three of the four mammalian Argonaute proteins (Broderick et al., 2011), quantitative mass spectrometry detected in the eluate only mouse AGO2 (Table S2.3), which was 64-fold enriched over the background control (Figure S2.2C). Perhaps the endogenous mouse AGO1, AGO3, and AGO4 proteins in the S100 are preoccupied with endogenous miRNAs and therefore unavailable for loading with exogenous siRNA.

Separation of a mixture of two RISCs

As a test of specificity, we asked whether our method can resolve RISCs programmed with different siRNAs. We separately assembled *let-7* and luciferase siRNAs into RISC using fly embryo lysate or mouse S100 cell extract. Next, we mixed the assembly reactions and then used a *let-7*-specific capture oligo to purify RISC from one half of the mixture. We used a luciferase-specific capture oligo to purify luciferase-siRNA-programmed RISC from the other half. Finally, we tested the eluates for their ability to cleave either *let-7* or luciferase target RNAs. RISC purified using the *let-7*-specific capture oligo only cleaved the *let-7* target (Figures 2.5 and 2.6). Conversely, RISC purified using the luciferase-specific capture oligo only cleaved the luciferase target. Thus, the method can separate a mixture of RISCs programmed with different small RNA sequences.

Figure 2.5

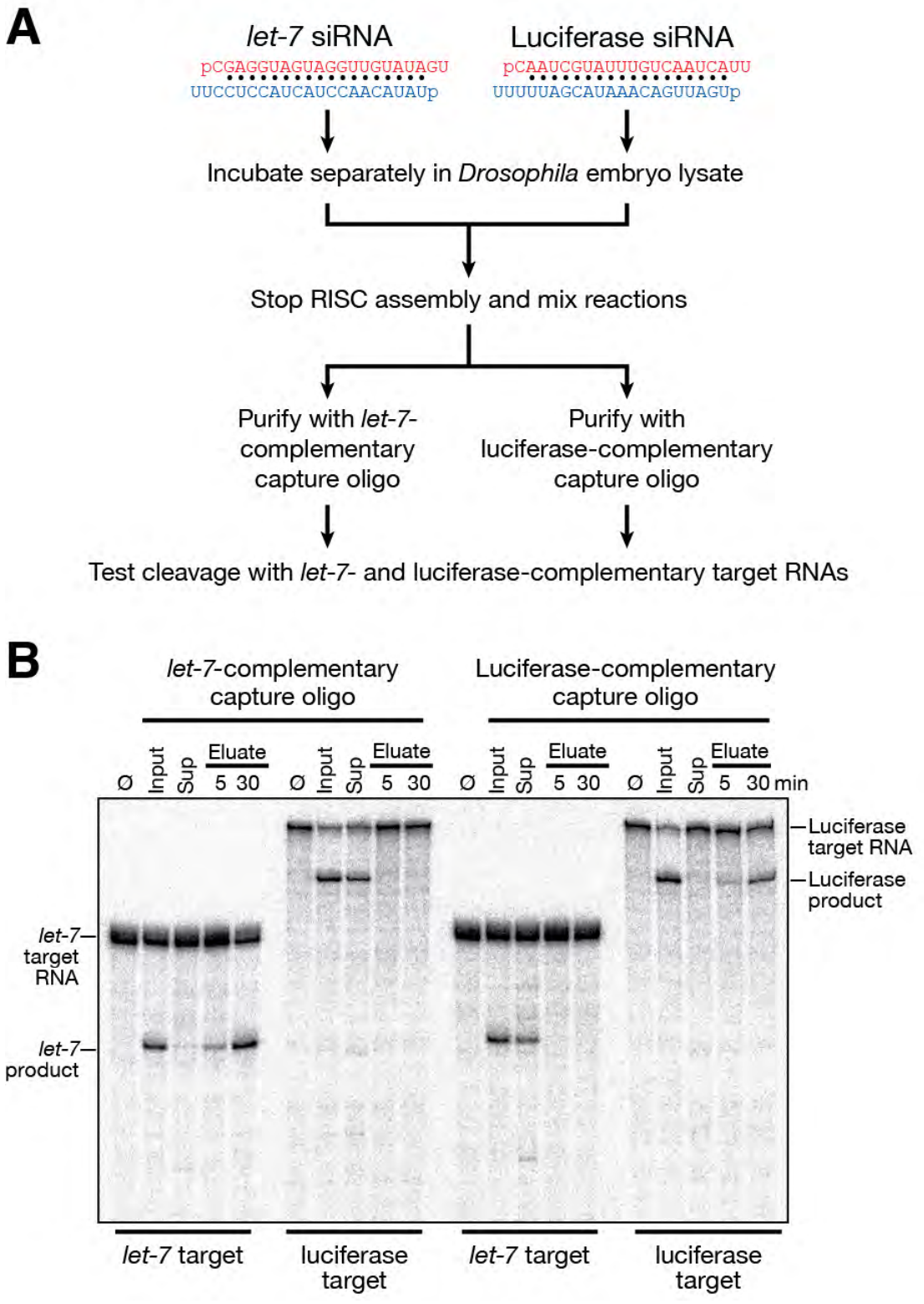

A
let-7 siRNA
pCGAGGUAGUAGGUUGUAUAGU
UUCCUCCAUCAUCCAACAUAUp
Luciferase siRNA
pCAAUCGUAUUUGUCAAUCAUU
UUUUUAGCAUAAACAGUUAGUp
Incubate separately in Drosophila embryo lysate
Stop RISC assembly and mix reactions
Purify with let-7-complementary capture oligo
Purify with luciferase-complementary capture oligo
Test cleavage with let-7- and luciferase-complementary target RNAs
B
let-7-complementary capture oligo
Luciferase-complementary capture oligo
Ø Input Sup Eluate 5 30 min
Luciferase target RNA
Luciferase product
let-7 target RNA
let-7 product
let-7 target
luciferase target
let-7 target
luciferase target

Figure 2.5: The purification method separates fly RISCs programmed with different siRNA guides.

(A) Experimental strategy. Red, guide; blue, passenger strand. (B) The activity of the samples was measured by incubating them with target RNAs (100 nM) complementary to let-7 (186 nt) and luciferase (506 nt) guide siRNA strands for 5 min (Input, Sup, Eluate) or 30 min (Eluate) at 25°C. Ø, no incubation.

Figure 2.6

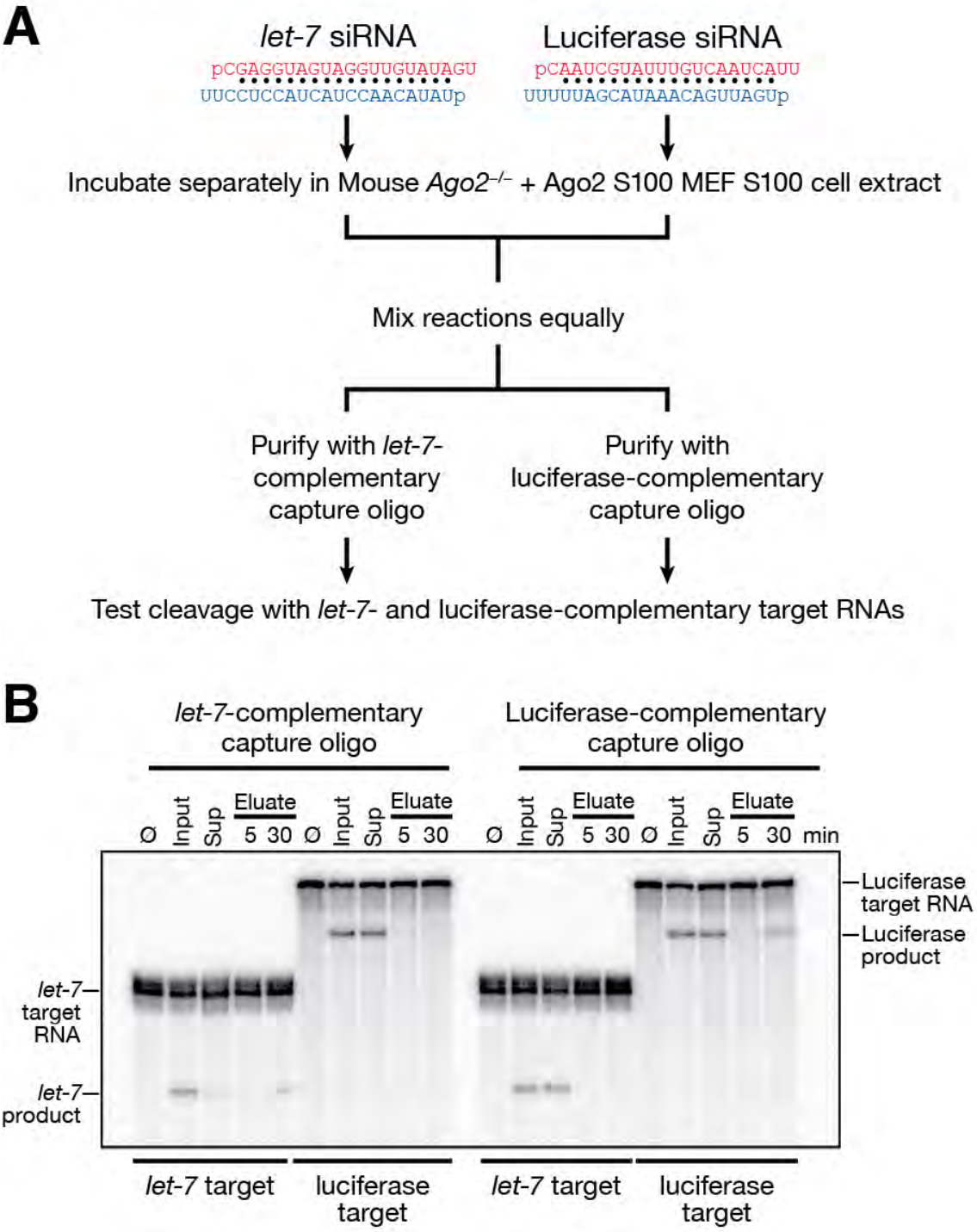

A
let-7 siRNA
pCGAGGUAGUAGGUUGUAUAGU
UUCCUCCAUCAUCCAACAUAUp
Luciferase siRNA
pCAAUCGUAUUUGUCAAUCAUU
UUUUUAGCAUAAACAGUUAGUp
Incubate separately in Mouse Ago2−/− + Ago2 S100 MEF S100 cell extract
Mix reactions equally
Purify with let-7-complementary capture oligo
Purify with luciferase-complementary capture oligo
Test cleavage with let-7- and luciferase-complementary target RNAs
B
let-7-complementary capture oligo
Luciferase-complementary capture oligo
Ø Input Sup Eluate 5 30 Ø Input Sup Eluate 5 30 Ø Input Sup Eluate 5 30 Ø Input Sup Eluate 5 30 min
Luciferase target RNA
Luciferase product
let-7—target RNA
let-7—product
let-7 target
luciferase target
let-7 target
luciferase target

Figure 2.6: The purification method separates mouse RISCs programmed with different siRNA guides.

(A) Experimental strategy. Red, guide; blue, passenger strand. (B) The activity of the samples was measured by incubating them with target RNAs (100 nM) complementary to let-7 (186 nt) and luciferase (506 nt) guide siRNA strands for 5 min (Input, Sup, Eluate) or 30 min (Eluate) at 37°C. Ø, no incubation.

Efficient purification with good yield

Both fully complementary and partially complementary capture oligos depleted all (Figures 2.2, 2.3A and 2.3B) or most (Figure 2.4D) detectable RISC activity from the lysate, but RISC was recovered efficiently only from the partially complementary capture oligo. To estimate the yield for our purification method, we measured both the amount of active purified RISC recovered in the eluate by pre-steady-state kinetic analysis and the differential loss of the passenger and guide strands (Figures S2.4). Nearly all the active RISC assembled was recovered from the partially complementary capture oligo: 78% for the capture oligo complementary to the seed plus the 3′ supplementary region and 80% for the capture oligo complementary to the seed alone (Table 2.1). Overall, the specific activity (pmole active RISC/mg protein) of fly Ago2-RISC increased 790-fold between the assembly reaction and the eluate of the capture oligo (Table 2.2). For mouse AGO2-RISC, the specific activity increased 320-fold.

Table 2.2

Species	Starting material			Final purified RISC			Purification factor
	Total protein (mg)	Active RISC (pmole)	Specific activity (pmole/mg)	Total protein (mg)	Active RISC (pmole)	Specific activity (pmole/mg)	
Fly	28	12	0.43	0.025	8.6	340	790
Mouse	3.3	2.6	0.78	0.0089	2.3	250	320

Table 2.2: Purification of *let-7* programmed Ago2-RISC from *Drosophila* embryo lysate and AGO2-over-expressing MEF S100

Purification of endogenously loaded miRNAs

We tested whether our method can also be used to purify RISC programmed with endogenous miRNAs. miR-21 is the most abundant miRNA in MEFs (J.A. Broderick and PDZ, unpublished data). We used a capture oligo complementary to the miR-21 seed plus 3′ supplementary region to purify miR-21-RISC from MEF S100 lysate, and then tested its ability to cleave a fully complementary miR-21 target RNA. Active miR-21-programmed RISC was recovered in the eluate (Figure 2.7A).

miR-286 is one of the most abundant miRNAs in 2–6 h *Drosophila* embryos (Ruby et al., 2007; Chung et al., 2008). We used a capture oligo complementary to the miR-286 seed plus 3′ supplementary region to purify miR-286-RISC from embryo lysate. Northern hybridization demonstrated that we could successfully recover miR-286 in the eluate of the capture oligo (Figure 2.7B).

Unlike purification of exogenously programmed RISC from fly embryo or mouse cell lysate, the capture oligo did not fully deplete the targeted miRNA from the lysate for either of the endogenous miRNAs. The failure to deplete miR-21- or miR-286-RISC from the lysate cannot be explained by an insufficient amount of capture oligo, because a second round of incubation with fresh capture oligo failed to deplete the remaining *Drosophila* miR-286 (Figure 2.7B). The simplest explanation is that as much as half of the miR-286 in the embryo

lysate is bound to its mRNA targets and is therefore unavailable to bind the capture oligo.

Figure 2.7

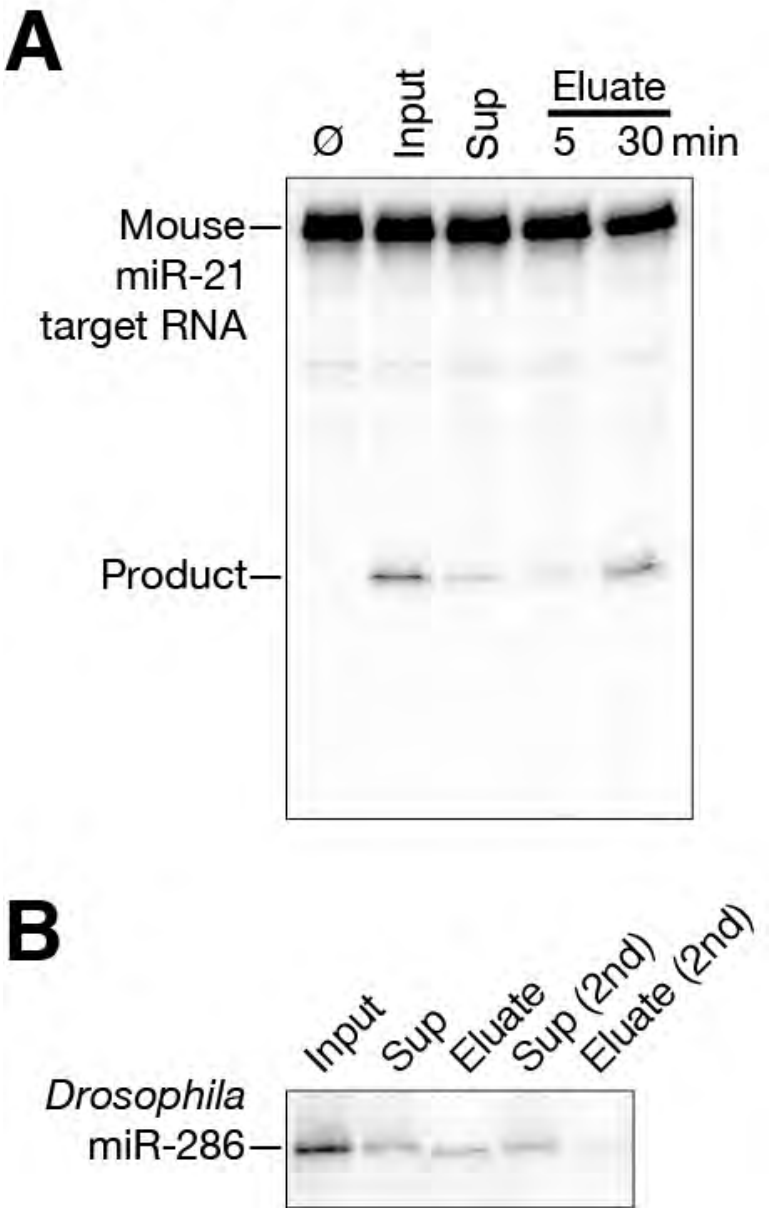

A
Ø
Input
Sup
Eluate
5
30 min
Mouse—
miR-21
target RNA
Product—
B
Input
Sup
Eluate
Sup (2nd)
Eluate (2nd)
Drosophila
miR-286—

Figure 2.7: Purification of endogenous miRNA-RISC complexes

(A) S100 cytosolic extract from *Ago2*$^{-/-}$ MEFs over-expressing mouse AGO2 was concentrated tenfold, and then purified using a capture oligo partially complementary to mouse miR-21. Target cleavage assays were performed by incubating samples with 100 nM target RNA for 5 min (Input, Sup, Eluate) or 30 min (Eluate) at 25°C. Ø, no incubation. (B) *Drosophila* miR-286 was purified from embryo lysate using a partially complementary capture oligo. Total RNA from each sample was resolved by denaturing electrophoresis, and miR-286 detected by northern hybridization.

DISCUSSION

We have described a simple method to purify from crude extracts active RISC complexes containing a single guide RNA sequence. Additionally, we describe a more robust cell extract system for the study of mammalian RNAi by over expressing AGO2 in MEF cells. Our procedure preserves the activity of target-cleaving Argonaute protein, but should also find use purifying Argonaute:small RNA complexes that do not catalyze endonucleolytic cleavage of their target RNAs.

Drosophila embryo lysate has been used extensively to study the biochemical mechanism of RNAi due to its robustness whereas there are fewer studies using mammalian cell extracts (Hutvagner and Zamore, 2002; Schwarz et al., 2002; Haley et al., 2003). A benefit to the use of cell extracts or lysates is that Argonaute proteins are programmed with their natural substrate, a siRNA duplex rather than using a single stranded RNA and recombinant protein (Rivas et al., 2005; MacRae et al., 2008). Why is embryo lysate more robust than mammalian extract system? There may be more Argonaute present in embryo lysate, however, a likely scenario is that the mammalian AGO proteins are already bound to miRNAs thus with no new AGO being translated in the cell extract, the exogenous siRNA cannot associate with AGO and does not form RISC. Alternatively, Drosophila reserves the AGO2 protein for viral siRNAs that are generated by Dicer upon infection (Wang et al., 2006). In our system, the

non-infected embryo lysate contains many unloaded AGO2 that is programmable (Haley et al., 2003). We used a MEF cell line that overexpresses AGO2 along with normal levels of other AGO proteins (O'Carroll et al., 2007; Broderick et al., 2011). This overexpression of AGO2 is at levels that satisfy the steady-state level of miRNAs while leaving an appreciable amount unloaded AGO that that is competent for siRNA assembly. This allows for the preparation of cell extracts that are reasonable in volume and concentration and, can be used to generate AGO2-RISC at concentrations that is suitable for in vitro biochemical studies.

Our method offers the advantage of purifying freely diffusing RISC complexes that can be used to study small RNA-directed Argonaute enzyme kinetics (Wee et al., 2012). Free diffusion is essential for quantitative population or single-molecule analysis of target binding by RISC.

Our method isolates only fully assembled, mature RISC containing a single-stranded RNA guide (Table 2.1). Thus, the proteins that co-purify with RISC should be only those present at the encounter with target. Our mass spectrometry analyses support this view, as we did not find proteins that participate in small RNA biogenesis or RISC assembly. We detected no peptides from Dicer-1, Dicer-2, R2D2, Loquacious, or C3PO (Jiang et al., 2005; Liu et al., 2006; Liu et al., 2007; Liu et al., 2009) co-purifying with *Drosophila* Ago1- or Ago2-RISC, nor did we detect Dicer, TRBP, or PACT (Hutvágner et al., 2001;

Chendrimada et al., 2005; Haase et al., 2005; Lee et al., 2006) co-purifying with mouse AGO2-RISC (Figure S2.3 and Tables S2.1–S2.3).

Heat shock proteins such as Hsp90 and Hsc70 also function in RISC assembly (Iki et al., 2010; Iwasaki et al., 2010; Johnston et al., 2010; Miyoshi et al., 2010; Miyoshi et al., 2010). Hsc70-4 was enriched over background for *Drosophila* Ago1-RISC purified by our method (Tables S2.1–S2.3). Hsc70-4 associates with Argonaute during its assembly into RISC, but Hsc70 and Hsp90 inhibitors had no detectable affect on target cleavage by Ago2 (Iwasaki et al., 2010); our data suggest Hsc70-4 may nonetheless play a role for Ago1 in target finding, target cleavage, product release, or enzyme regeneration after target cleavage. Whereas Ago2 was the most enriched protein in the Ago2 sample, other proteins were highly enriched in the Ago1 sample (Table S2.1); ribosomal protein L7A was the only protein enriched in both samples. We do not yet know if these Argonaute-associated proteins play a role in RISC function.

Although the siRNAs and miRNAs used here contained no chemical modifications, we have used the method to purify RISC programmed with fluorescently labeled siRNA guide strands (CFF-J and PDZ, unpublished data). We envision that our method will prove useful in the biochemical dissection of the RNAi and miRNA pathways and in the identification of proteins that modify and extend RISC function.

Materials and Methods

General methods

Synthetic DNA (IDT, Coralville, IA) and RNA (GE Healthcare Dharmacon, Lafayette, CO) oligonucleotides were gel purified before use. RNA oligonucleotides were synthesized containing a 5′ phosphate. The strands of the let-7 and luciferase siRNA duplexes were annealed in lysis buffer (30 mM HEPES-KOH, pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate) using a 1:1.3 molar ratio of guide to passenger. The let-7/let-7* duplex was annealed using a 1:5 molar ratio of miRNA to miRNA*.

Drosophila embryo lysate and MEF S100 cytosolic extract Drosophila embryo lysate from 0–8 h embryos and RISC assembly were as described (Haley et al., 2003). The protein concentration of the embryo lysate was ~30 mg/ml.

S100 extract was generated from SV40 large T-antigen immortalized *Ago2*$^{-/-}$ MEFs that stably over-express mouse Ago2 mRNA (O'Carroll et al., 2007), SV40 large-T antigen immortalized MEF (Salomon and Zamore, unpublished) or HeLa cells (ATCC, Manassas, VA). Cells were grown to confluence in 5% CO2 at 37°C in DMEM (Thermo Fisher Scientific – Life Technologies, Grand Island, NY) supplemented with 15% fetal bovine serum (GE Healthcare Bio-Sciences, PAA Laboratories, Pittsburg, PA) and 50 U/ml

penicillin and streptomycin (Thermo Fisher Scientific – Life Technologies). S100 extract was prepared as described (Dignam et al., 1983) except that the cell pellet was washed three times in ice-cold PBS and once in Buffer A (10 mM HEPES–KOH, pH 7.9, 10 mM potassium acetate, 1.5 mM magnesium acetate, 0.5 mM DTT, EDTA-free protease inhibitor cocktail). The supernatant was removed, and 0.11 cell pellet volumes of Buffer B (300 mM HEPES-KOH, pH 7.9, 1.4 M potassium acetate, 30 mM magnesium acetate, 0.5 mM DTT, EDTA-free protease inhibitor cocktail) was added, followed by centrifugation at 100,000 × g at 4°C for 20 min. Ice-cold 80% (w/v) glycerol was then added to achieve a 13% (w/v) final glycerol concentration, followed by gentle inversion to mix. S100 was aliquoted, frozen in liquid nitrogen, and stored at –80°C. S100 protein concentration was ~3–4 mg/ml.

Capture oligonucleotides and siRNAs

Streptavidin paramagnetic beads (100 µl; Dynabeads MyOne Streptavidin T1 10 mg/ml, Thermo Fisher Scientific – Life Technologies) were washed and incubated with 5′ biotinylated, 2′-*O*-methyl capture oligonucleotides (500 pmol) according to the manufacturer's instructions, then re-suspended in 100 µl of lysis buffer containing 2 mM dithiothreitol (DTT) and kept on ice for RISC purification. Capture oligonucleotides (entirely 2′-*O*-methyl ribose) were: 5′-biotin-UCU UCA CUA UAC AAC CUA CUA CCU CAA CCU U-3′ (fully complementary to let-7), 5′-biotin-UCU UCC UGC GAC AAU AGC CUA CCU

CAA CCU U-3′ (seed plus 3′ supplementary pairing for let-7); 5′-biotin-UCU UCC UGC GCA CCA AGC CUA CCU CAA CCU U-3′ (seed pairing to let-7); 5′-biotin-UCU GAC GCA CUU GAU UCU UAC GAU UUA UCU A-3′ (seed plus 3′ supplementary pairing for luciferase siRNA guide); 5′-biotin-GAU GAA CCA CUC AGA GAC AUA AGC UAA UCU A-3′ (seed plus 3′ supplementary pairing for mmus-miR-21); 5′-biotin-UCU GAC AAC GUU GUG UAA CCU CUA GUC CAU CU-3′ (seed plus 3′ supplementary pairing for Drosophila miR-286). siRNA sequences were 5′-pCGA GGU AGU AGG UUG UAU AGU-3′ (used as let-7 siRNA guide in Drosophila embryo lysate) or 5′-pUGA GGU AGU AGG UUG UAU AGU-3′ (used as let-7 siRNA guide in MEF lysate); 5′-pUAU ACA ACC UAC UAC CUC CUU-3′ (let-7 passenger); 5′-pCAA UCG UAU UUG UCA AUC AGA-3′ (luciferase guide); 5′-pUGA UUG ACA AAU ACG AUU UUU-3′ (luciferase passenger). miRNA sequences were 5′-pUGA GGU AGU AGG UUG UAU AGU-3′ (let-7); 5′-pUAU ACA AUG UGC UAG CUU UCU-3′ (let-7*).

RISC purification

The buffer from the washed capture oligo paramagnetic beads (500 pmole capture oligo on 1 mg beads) was removed and replaced with 100 µl RISC assembly reaction, then incubated with gentle rotation at room temperature for 30 min. The supernatant was removed, and the beads were washed five times with lysis buffer containing 2 mM DTT and 0.01% (v/v) Triton X-100, followed by five washes with lysis buffer containing 2 M potassium acetate, DTT, and Triton

X-100. RISC was eluted by re-suspending the beads in 100 µl of lysis buffer containing 1 M potassium acetate, DTT, Triton X-100, and 1 nmol (10 µM f.c) competitor oligo and incubating with gentle rotation for 120 min at room temperature. Competitor sequences were 5′-AAG GTT GAG GTA GTA GGT TGT ATA GTG AAG A-3′ (for the capture oligo fully complementary to let-7); 5′-AAG GTT GAG GTA GGC TAT TGT CGC AGG AAG A-3′ (for the capture oligo with seed plus 3′ supplementary pairing to let-7); 5′-AAG GTT GAG GTA GGC TTG GTG CGC AGG AAG A-3′ (for seed pairing to let-7); 5′-TAG ATA AAT CGT AAG AAT CAA GTG CGT CAG A-3′ (for seed plus 3′ supplementary to luciferase). 5′-AGA TGG ACT AGA GGT TAC ACA ACG TTG TCA GA-3′, (for seed plus 3′ supplementary pairing to Drosophila miR-286); 5′-UAG ATT AGC UTA UGT CTC TGA GUG GTT CAT C-3′, (for seed plus 3′ supplementary pairing to mouse miR-21). Finally, all samples were dialyzed at 4°C against three changes (2 h each) of a 2,000-fold excess of lysis buffer containing 2 mM DTT.
For quantitative analysis of yield (Table 2.1), guide strands were 5′ ^{32}P-radiolabeled.

Ion exchange chromatography on Mono S 5/50 GL (GE Healthcare) was used to remove competitor DNA. The column was equilibrated in lysis buffer containing 2 mM DTT. After loading, the column was washed with five column volumes lysis buffer containing 2 mM DTT. RISC was eluted with lysis buffer containing 2 mM DTT and 500 mM potassium acetate. The competitor DNA

eluted in the wash. When necessary, samples were concentrated by centrifugal ultrafiltration (10 kDa cutoff; EMD Millipore, Billerica, MA).

RISC activity and concentration

Target cleavage assays were as described (Haley et al., 2003). Target RNAs were transcribed with T7 RNA polymerase from templates generated by PCR of pGL2 plasmid (Promega, Madison, WI) using these primers: 5′-GCG TAA TAC GAC TCA CTA TAG GGT CAC ATC TCA TCT ACC TCC-3′ (let-7a forward); 5′-CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G-3′ (let-7a reverse); 5′-GAT GCG TAA TAC GAC TCA CTA TAG GGT TCC GCA TAG AAC TGC CTG CGT CA-3′ (luciferase forward); 5′-TCC AGA TCC ACA ACC TTC GCT TCA-3′ (luciferase reverse); 5′-GCG TAA TAC GAC TCA CTA TAG GGG TCC TTT GAT CGT GAC AAA ACA AT-3′ (mouse miR-21 forward); 5′-CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC TAG TTG AGG TGC GGA ACT GTG TAT AAA AGG TTA GCT TAT CAG ACT GAT GTT GAA TCC AGA GGA ATT CAT TAT CAG TG-3′ (mouse miR-21 reverse).

In Figure S2.4, reaction products were resolved by denaturing polyacrylamide gel electrophoresis, and the amount of cleaved target was quantified using an FLA-9000 phosphorimager (GE Healthcare) and ImageGauge software (Fuji Life Sciences, Tokyo). IC50 values were obtained by fitting (Igor Pro 6.22, WaveMetrics, Oswego, OR) the data to the Hill equation assuming non-

cooperative behavior. Pre-steady-state kinetic data were fit in Igor Pro to the burst-and-steady-state equation,

$$F(t) = E \times \frac{a^2}{(a+b)^2}\left(1 - e^{-(a+b)t}\right) + E \times \frac{ab}{(a+b)}t$$

where F(t) is target cleaved with time, E is the enzyme concentration, and a and b are rate constants according to the following scheme,

$$E + S \xrightarrow{a} ES \xrightarrow{b} E + P$$

The concentration of RISC assembled in Drosophila embryo lysate (Figure S2.4) was estimated by measuring the loss of 5′ ^{32}P-radiolabeled passenger and guide strands. Control reactions, in which the lysate was pre-incubated with 1 mM (f.c.) *N*-ethylmaleimide (NEM) at 4°C for 10 min to prevent RISC assembly (Nykanen et al., 2001), demonstrated that differential loss of passenger and guide strands required RISC assembly. Unreacted NEM was quenched with 1.2 mM (f.c.) DTT. Reaction products were resolved by denaturing 20% polyacrylamide gel electrophoresis and detected as described. The rate of siRNA decay was determined separately for each replicate, and the significance determined using Student's two-sample, t-test (GraphPad Prism 5, GraphPad Software).

Western Blotting

S100 cell extract was diluted in lysis buffer containing 1M DTT to 2.5 mg·ml total protein concentration. The diluted S100 extract was mixed 1:1 in 2× Laemmli sample buffer (120 mM Tris-HCl, pH 6.8, 10% (w/v) SDS, 20% v/v glycerol, 0.02% (w/v) bromophenol blue, 100 mM fresh DTT) and boiled at 95°C for 5 minutes. Samples were separated by 4-20% HEPES-SDS-PAGE and transferred at 4°C in Tris-glycine buffer to nitrocellulose membrane overnight at 30 V. The membrane was blocked in fluorescent western blot blocking buffer (Rockland, Limerick, PA) diluted 1:1 in TBS-T buffer for 1 h and incubated overnight at 4°C with primary antibody diluted in fluorescent western blot blocking buffer (Rockland) diluted 1:1 in TBS-T. Rabbit anti-human and mouse Ago2 antibody (Cell Signaling Technologies, Danvers, MA) (Li et al., 2010) was diluted 1:1,000 and mouse anti-β-tubulin antibody (Sigma, St. Louis, MO) was diluted 1:3,000. After three 5 min washes in TBST the membranes were incubated for 1 h at room temperature in the dark with secondary goat anti-rabbit 800CW IRDye-conjugated antibody (Li-Cor, Lincoln, NE) and secondary goat anti-mouse 680RD IRDye-conjugated antibody (Li-Cor), both diluted 1:10,000. After three, 15 min washes in TBST, the membrane was placed on an Odyssey Imager where 785 nm and 685 nm lasers excited the secondary antibody. Images were quantified using Image Studio (Li-Cor).

Northern hybridization

Total RNA (5 µg), isolated from 2–6 h fly embryo lysates using Trizol (Thermo Fisher Scientific – Life Technologies) according to manufacturer's instructions, was resolved on a 0.4 mm thick, 15% denaturing polyacrylamide sequencing gel. After electrophoresis, the gels were blotted at 20 V for 1 h to a Hybond-NX membrane (GE Healthcare) in 0.5× TBE (Tris-Borate-EDTA) buffer by semi-dry transfer (Trans-Blot SD, Bio-Rad, Hercules, CA). The RNA was cross-linked to the membrane with freshly prepared 140 mM l-ethyl-3-(3-dimethylaminopropyl) carbodiimide (Sigma) in 130 mM 1-methylimidazole (Sigma), pH 8, for 1 h at 60°C. The DNA oligonucleotide probe, Drosophila miR-286 5′-AGCACGAGTGTTCGGTCTAGTCA-3′ (25 pmol), was 5′-radiolabeled with T4 polynucleotide kinase (New England Biolabs, Ipswich, MA) and γ-^{32}P-ATP (450 µCi per reaction, specific activity ~6000 Ci/mmol; Perkin Elmer, Waltham, MA). After labeling, unincorporated nucleotides were removed (Sephadex G-25 spin column, GE Healthcare), and the probes were added to the Church buffer (Church, 1984) and hybridized overnight at 37°C. Membranes were washed three times for 20 min in 0.01× SSC containing 0.1% SDS (w/v) at 37°C and exposed to storage phosphor screens (GE Healthcare).

Mass spectrometry

For mass spectrometry, RISC was purified from a 2 ml assembly reaction and then concentrated tenfold. Proteins were separated from low molecular weight contaminants by a brief period of SDS polyacrylamide gel electrophoresis, then a single gel slice containing all protein was excised, eluted and digested with trypsin. For comparison, control reactions in which the small RNA duplex was omitted were included (Figures 2.3C and S2.3, background and minus siRNA controls).

Tryptic peptides were dissolved in 0.1% (v/v) trifluoroacetic acid, loaded onto a fused silica trap column (2 cm × 100 μm C18), and then fractionated on a fused silica column (25 cm × 75 μm C18) developed with a linear gradient from 100% solvent A (0.1% formic acid, 5% acetonitrile in water) to 35% solvent B (0.1% formic acid in water) at a flow rate of 300 nl min-1 over 90 min using a nano LC: EASY system (Thermo Scientific – Proxeon) directly coupled to an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific). Data-dependent acquisitions from MS scans (350–2000 m/z) in the Orbitrap at resolution, r = 60,000 were followed by ten MS/MS scans acquired in the LTQ ion trap instrument.

Raw data files were processed with Mascot Distiller (Matrix Science, Boston) or Extract-MSn (Thermo Scientific) to generate peak lists and then searched against the Drosophila melanogaster (NCBInr) or Mus musculus

(UniProtKB/Swiss-Prot) index using Mascot Search engine 2.3.02 (Matrix Science) with 10 ppm parent mass and 0.5 Da fragment mass tolerances. Amino acid modifications considered were acetyl (for the amino-terminus of the protein), pyroglutamic acid (for amino-terminal glutamine), propionamide and carbamidomethylation modification of cysteine, and oxidation of methionine. Label-free quantification using extracted ion chromatograms (XIC) was performed using both the replicate and average methods in the Mascot Distiller quantitation software (Matrix Science). Mascot searches were also loaded into Scaffold3 software (Proteome Software, Portland, OR) for further comparative analyses and filtering.

ACKNOWLEDGEMENTS

We thank Dónal O'Carroll for *Ago2*$^{-/-}$ MEFs over-expressing mouse AGO2; John Leszyk and the UMass Medical School Proteomics and Mass Spectrometry Facility for mass spectrometry; Alicia Boucher and Cindy Tipping for fly husbandry; and members of the Zamore laboratory for help, discussions, advice, and comments on the manuscript. This work was supported in part by grants from the National Institutes of Health to PDZ (GM62862 and GM65236), and a CONACyT Postdoctoral Fellowship to CFF-J.

Figure S2.1

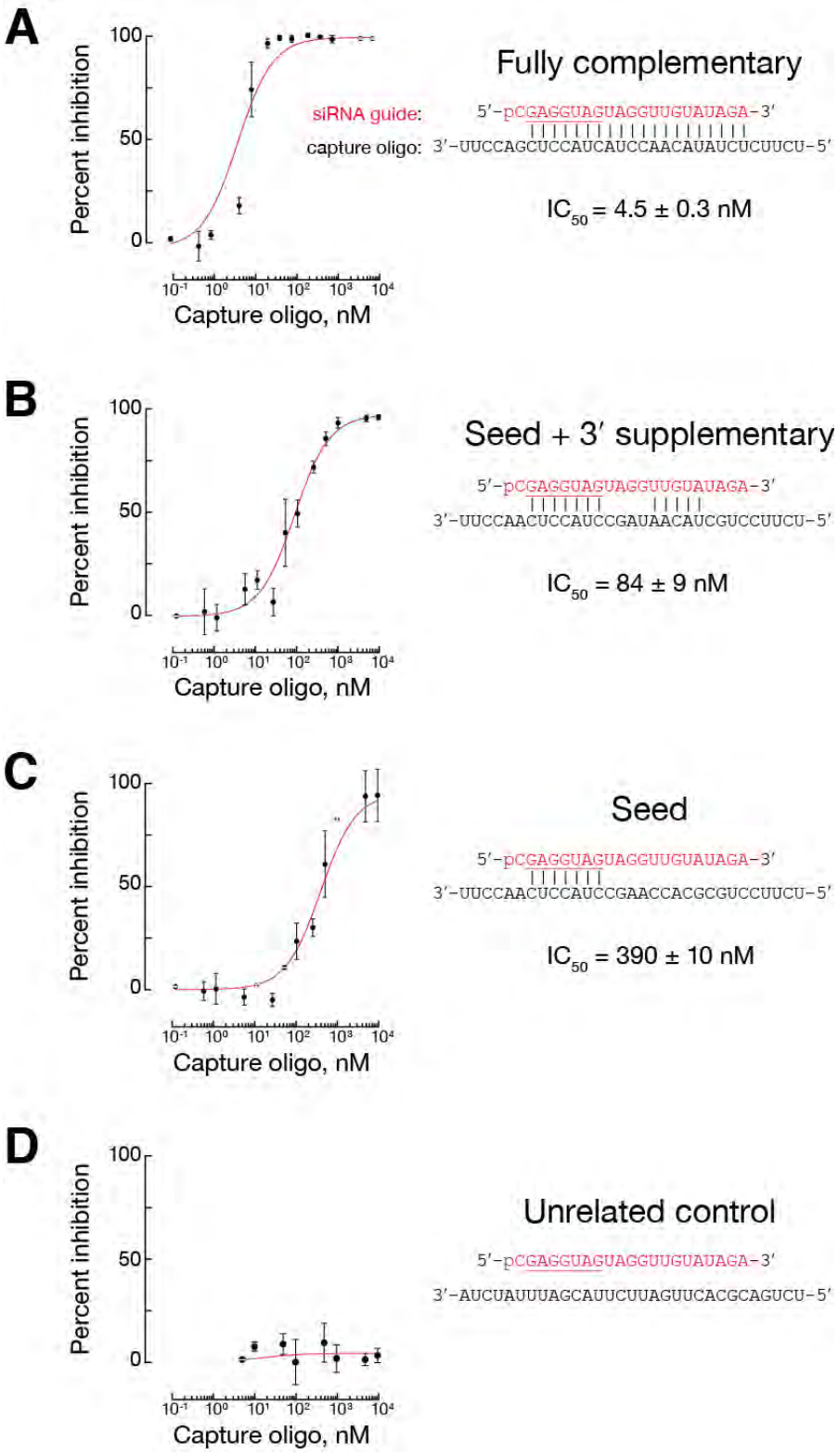

A
Percent inhibition
100
50
0
10⁻¹ 10⁰ 10¹ 10² 10³ 10⁴
Capture oligo, nM
Fully complementary
siRNA guide: 5′-pCGAGGUAGUAGGUUGUAUAGA-3′
capture oligo: 3′-UUCCAGCUCCAUCAUCCAACAUAUCUCUUCU-5′
IC50 = 4.5 ± 0.3 nM
B
Percent inhibition
100
50
0
10⁻¹ 10⁰ 10¹ 10² 10³ 10⁴
Capture oligo, nM
Seed + 3′ supplementary
5′-pCGAGGUAGUAGGUUGUAUAGA-3′
3′-UUCCAACUCCAUCCGAUAACAUCGUCCUUCU-5′
IC50 = 84 ± 9 nM
C
Percent inhibition
100
50
0
10⁻¹ 10⁰ 10¹ 10² 10³ 10⁴
Capture oligo, nM
Seed
5′-pCGAGGUAGUAGGUUGUAUAGA-3′
3′-UUCCAACUCCAUCCGAACCACGCGUCCUUCU-5′
IC50 = 390 ± 10 nM
D
Percent inhibition
100
50
0
10⁻¹ 10⁰ 10¹ 10² 10³ 10⁴
Capture oligo, nM
Unrelated control
5′-pCGAGGUAGUAGGUUGUAUAGA-3′
3′-AUCUAUUUAGCAUUCUUAGUUCACGCAGUCU-5′

Figure S2.1: Determining the concentration of fully and partially complementary 2′-*O*-methyl oligonucleotides required to inhibit RISC activity.

let-7 duplex siRNA (guide strand, red; passenger strand, blue) was assembled into Ago2-RISC in Drosophila embryo lysate. Then, 2′-*O*-methyl oligonucleotides (black) fully complementary to let-7 (A), complementary to the let-7 seed plus the 3′ supplementary region (B), to the let-7 seed alone (C), or not complementary to let-7 (D) were added and incubated at 25°C for 30 min. Finally, ^{32}P-radiolabeled target RNA was added (50 nM f.c.), incubated for 5 min, and target cleavage measured. Data are mean ± S.D., n = 3.

Figure S2.2

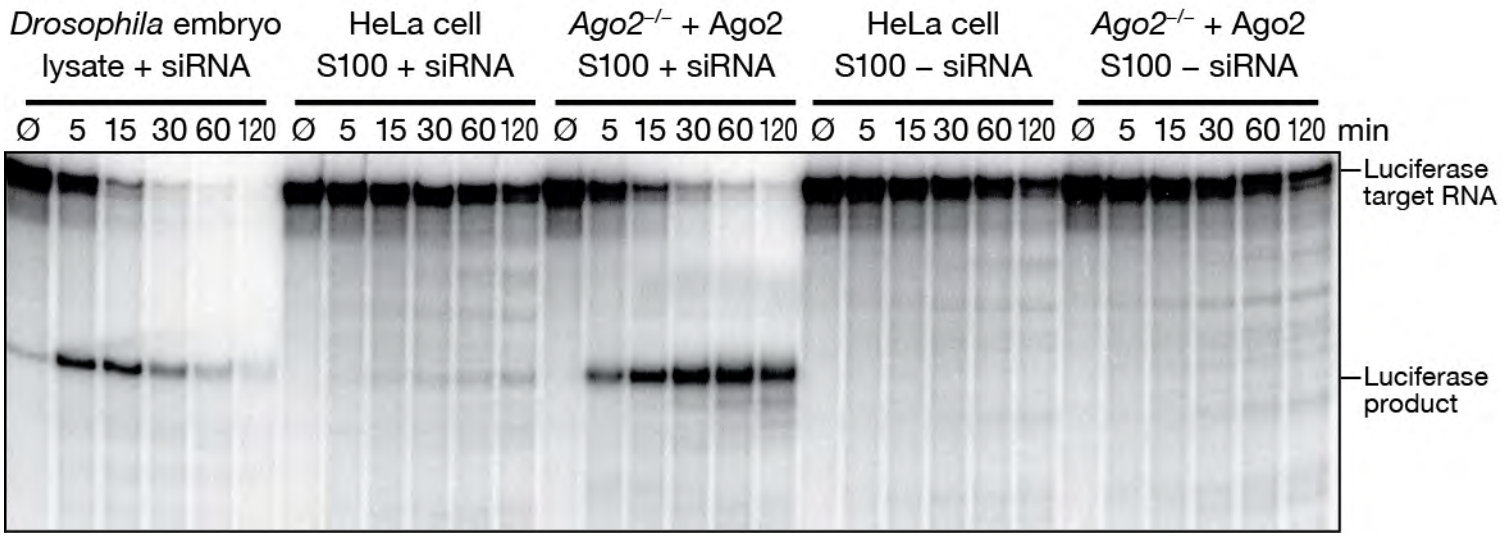

Drosophila embryo lysate + siRNA
HeLa cell S100 + siRNA
Ago2$^{-/-}$ + Ago2 S100 + siRNA
HeLa cell S100 – siRNA
Ago2$^{-/-}$ + Ago2 S100 – siRNA
Ø 5 15 30 60 120 Ø 5 15 30 60 120 Ø 5 15 30 60 120 Ø 5 15 30 60 120 Ø 5 15 30 60 120 min
Luciferase target RNA
Luciferase product

Figure S2.2: Comparison of HeLa cell and Ago2 overexpressing MEF cell S100 cytosolic extracts for RNAi activity

Drosophila embryo lysate was programmed with 25 nM luciferase siRNA at 25°C while HeLa and MEF S100 cell extracts were programmed with or without (buffer) 25 nM luciferase siRNA at 37°C. Target cleavage assay was carried out using 100 nM capped ^{32}P-labeled luciferase RNA target. Samples were resolved on a urea-denaturing polyacrylamide gel and the gel was dried and exposed as described in the Material and Methods.

Figure S2.3

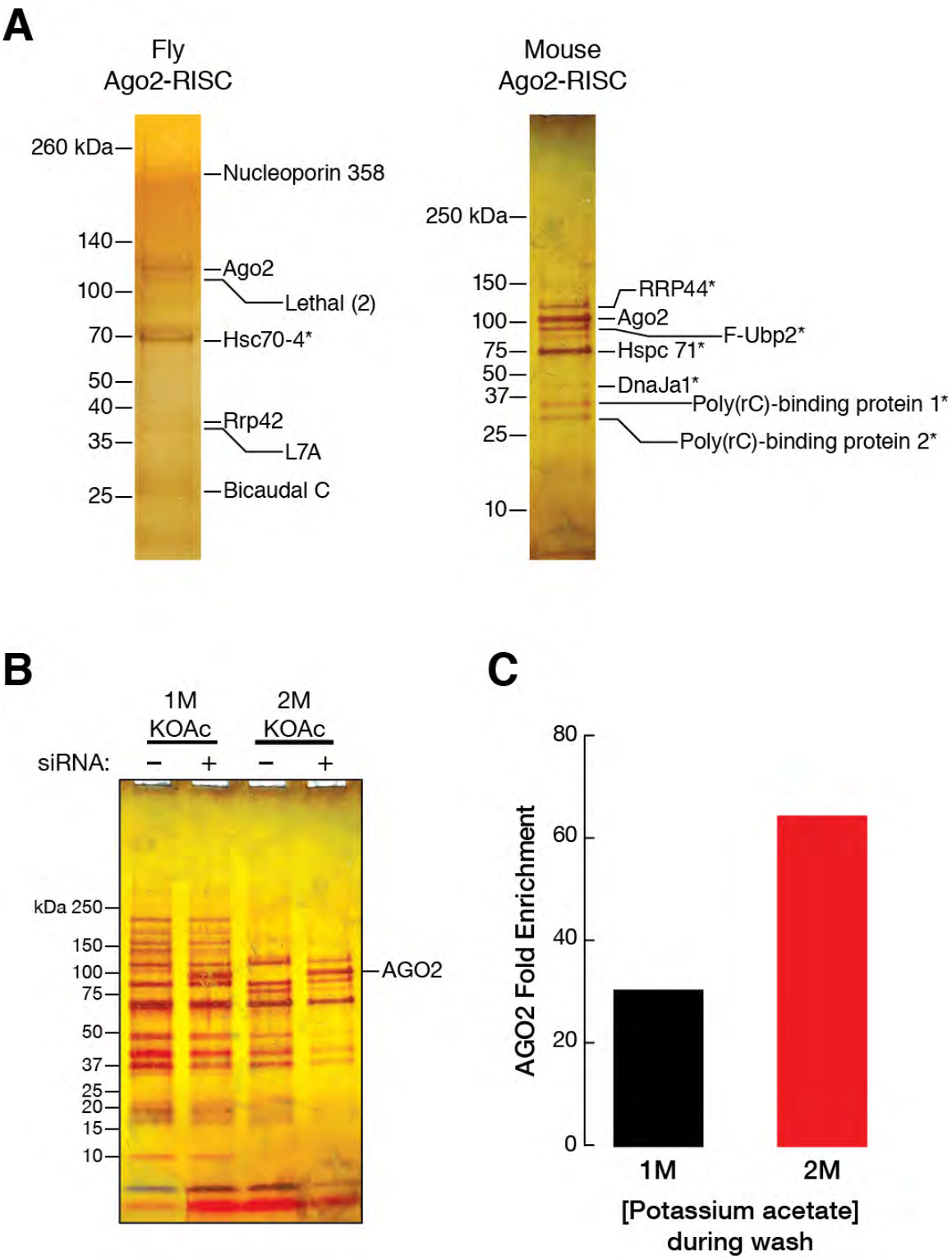

A
Fly
Ago2-RISC
260 kDa
140
100
70
50
40
35
25
Nucleoporin 358
Ago2
Lethal (2)
Hsc70-4*
Rrp42
L7A
Bicaudal C
Mouse
Ago2-RISC
250 kDa
150
100
75
50
37
25
10
RRP44*
Ago2
F-Ubp2*
Hspc 71*
DnaJa1*
Poly(rC)-binding protein 1*
Poly(rC)-binding protein 2*
B
1M KOAc
2M KOAc
siRNA: – + – +
kDa 250
150
100
75
50
37
25
20
15
10
AGO2
C
AGO2 Fold Enrichment
80
60
40
20
0
1M
2M
[Potassium acetate] during wash

Figure S2.3: Silver-stained 4–20% acrylamide gel loaded with 50 fmol affinity purified let-7–programmed fly or mouse Ago2-RISC

(A) Proteins whose identities were established by mass spectrometry are labeled. Asterisks denote proteins that were not enriched compared to control and are therefore likely to be contaminants. (B) Silver stained 4–20% SDS-PAG comparing 1M and 2M potassium acetate washes in RISC purification. MEF cells overexpression AGO2 were incubated with or without let-7 siRNA prior to purification. The observed bands for with or without siRNA programming are similar with the exception of AGO2. (D) Mass spectrometry comparison of AGO2 enrichment over no siRNA control. AGO2 was enriched 30-fold for 1M potassium acetate wash and 64-fold for the 2M potassium acetate wash.

Figure S2.4

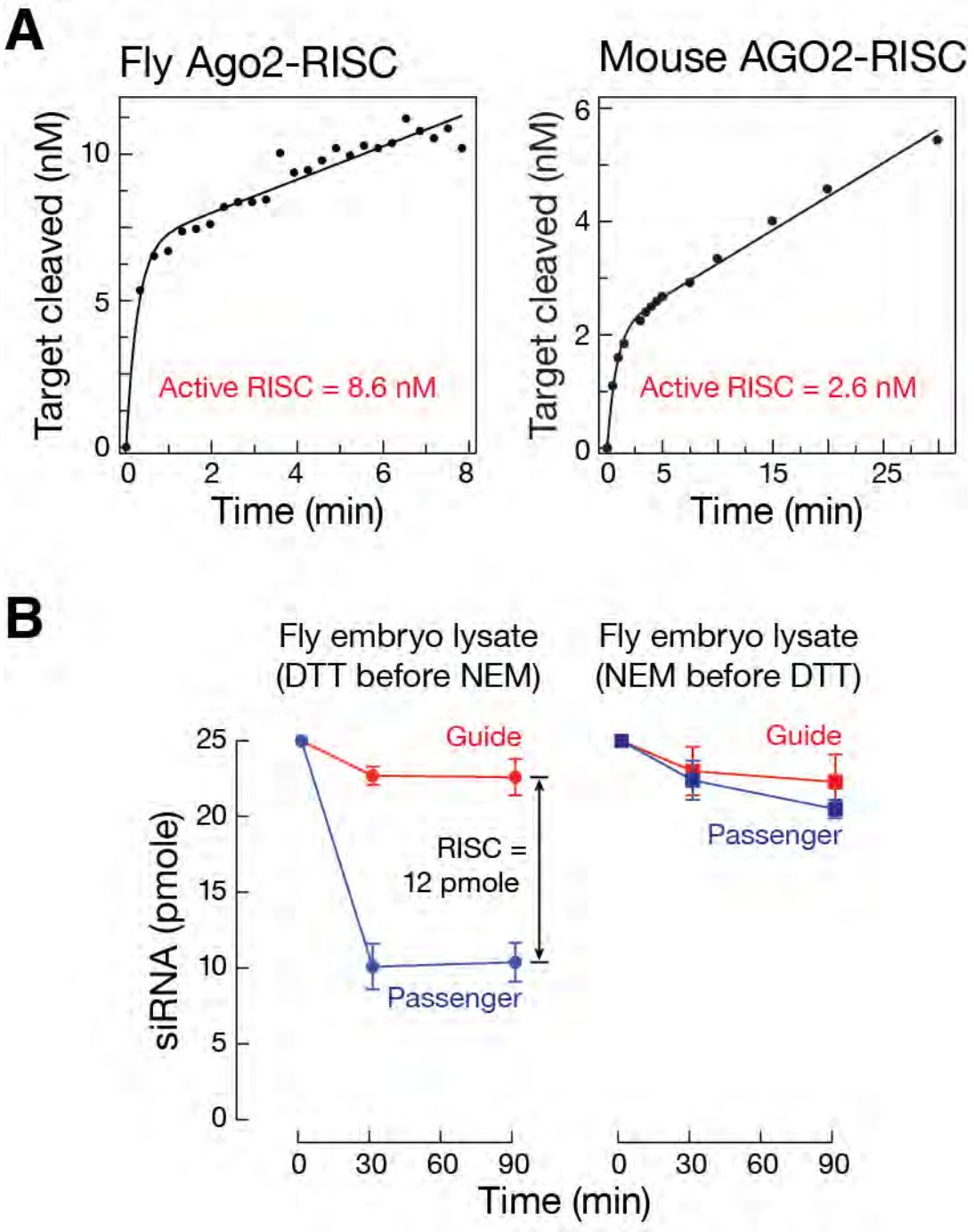

A
Fly Ago2-RISC
Target cleaved (nM)
Active RISC = 8.6 nM
Time (min)
Mouse AGO2-RISC
Target cleaved (nM)
Active RISC = 2.6 nM
Time (min)
B
Fly embryo lysate
(DTT before NEM)
Fly embryo lysate
(NEM before DTT)
siRNA (pmole)
Guide
Passenger
RISC =
12 pmole
Time (min)

Figure S2.4: Measuring RISC concentration

(A) The target cleaving activity of purified fly and mouse AGO2-RISC was measured by pre-steady state kinetics using 100 nM 5′ ^{32}P-radiolabeled target RNA. For fly RISC, the reaction was incubated at 10°C; for mouse RISC, the reaction was incubated at 23°C. Total enzyme concentration was determined by fitting the data to the burst-and-steady-state equation (see Methods).
(B) Drosophila embryo lysates were incubated with let-7a siRNA duplex (25 nM f.c.), 5′ ^{32}P-radiolabeled on the guide or passenger strand. RISC assembly reactions mock-treated with *N*-ethylmaleimide (NEM): dithiothreitol (DTT) was added before the NEM. The differential decay of the guide and passenger strands (12 pmole; p = 0.0001) indicates the amount of RISC assembled. In the control reactions, RISC assembly was blocked by treating with NEM. Unreacted NEM was subsequently quenched with DTT. Data are mean ± S.D., n = 3.

Table S2.1: Proteins enriched in purified fly Ago1-RISC compared to the control

Mass (kDa)	Ago1 / control	Protein
17603	3595	CG3800, isoform A
27875	1339	Ribosomal protein L8, isoform A
30713	1086	Ribosomal protein L7A, isoform D
71087	917	Heat shock protein cognate 4, isoform A
24258	851	Ribosomal protein L10Ab, isoform A
41004	639	Hrp48.1
29718	590	Ribosomal protein L6, isoform B
24020	589	Ribosomal protein L19
71051	549	CG7185
106150	467	**Argonaute1, isoform B**
69882	425	Poly(A)-binding protein, isoform A
63368	386	RE72930p
62091	310	IGF-II mRNA-binding protein, isoform A
25765	283	GM01970p
69298	277	Trailer hitch
12227	252	Ribosomal protein L30, isoform C
102887	229	CG8108, isoform B
40624	217	Mushroom-body expressed, isoform A
25010	207	Rrp40
74898	190	Rasputin, isoform B
50208	188	RE71384p
34616	185	Fibrillarin
46274	183	Rrp45
26933	153	RE67757p
103528	132	CG18811
18127	131	CG8928
102854	130	Rrp6
36219	119	Hrp40.2
15893	111	Ribosomal protein L27

112140	101	SD10981p
23058	93	Ribosomal protein S6, isoform B
29534	92	Ribosomal protein L7
37150	78	Replication factor C subunit 4
78000	61	Protein on ecdysone puffs, isoform B
25700	52	Ribosomal protein S5b
85029	40	Belle
116767	33	CG16940, isoform A
33589	28	GM02257p
32366	24	Rrp42
118850	14	Lethal (2) 35Df
75773	6	Fmr1, isoform C
24310	6	Ribosomal protein L15, isoform A
95704	5	Argonaute2
136765	4	Argonaute2, isoform B
105665	3	CG14476, isoform B
296201	3	Nucleoporin 358
57007	2	DNop5 protein
98496	2	Aubergine, isoform A
72190	2	Heat shock protein cognate 72
50760	1	Phosphoprotein phosphatase 2A 55 kDa regulatory subunit
41797	1	Actin

Table S2.2: Proteins enriched in purified fly Ago2-RISC compared to the control

Mass (kDa)	Ago2 control	Protein
95704	420	**Argonaute 2**
136765	291	**Argonaute 2, isoform B**
30713	229	Ribosomal protein L7A, isoform D
296201	138	Nucleoporin 358
98496	118	Aubergine, isoform A
32366	45	Rrp42
97788	25	Bicaudal C, isoform A
33589	4	GM02257p
112140	3	SD10981p
103528	3	CG18811
24020	3	Ribosomal protein L19
74898	3	Rasputin, isoform B
88155	3	CG10077, isoform A
48682	2	Yolk protein 1
62091	2	IGF-II mRNA-binding protein, isoform A
63368	2	RE72930p
69882	2	Poly(A)-binding protein, isoform A
30209	2	Stubarista, isoform A
41004	2	Hrp48.1
29718	2	Ribosomal protein L6, isoform B

106150	1	Argonaute1, isoform B
100356	1	CG10777
27875	1	Ribosomal protein L8, isoform A
17603	1	CG3800, isoform A
46274	1	Rrp45target cleaving
40624	1	Mushroom-body expressed, isoform A
118850	1	Lethal (2) 35Df

Table S2.3: Proteins enriched in purified mouse AGO2-RISC compared to the control

Mass (kDa)	AGO2 / Control	Protein
97	64	**Argonaute2**
50	1	Tubulin β-5 chain
72	1	78 kDa Glucose-regulated protein
50	1	Tubulin beta-2C chain
71	1	Heat shock cognate 71
77	1	Far upstream element-binding protein 2
109	1	Exosome complex exonuclease RRP44
70	<1	Splicing factor 1
103	1	RNA-binding protein 12
69	<1	Far upstream element-binding protein 1
70	1	Heat shock 70 kDa protein 1A

Chapter III: Single-Molecule Imaging Reveals that Argonaute Re-shapes the Binding Properties of Its Nucleic Acid Guides

Disclaimer

This work is the joint effort among the authors: William E. Salomon (WES), Samson M. Jolly (SMJ), Melissa J. Moore (MJM), Phillip D. Zamore (PDZ), and Victor Serebrov (VS). SMJ prepared and optimized small DNA assembly of *Thermus thermophilus* Argonaute. VS performed single-molecule microscopy, wrote custom Matlab scripts for analysis, and assisted with data analysis. WES prepared all reagents (with the exception of *Thermus thermophilus* AGO), performed all in vitro analyses, single-molecule analyses, and prepared all of the figures. WES and PDZ wrote the manuscript. All authors contributed to critical review of data and manuscript. Addition of unpublished and supporting data has been included in this chapter.

SUMMARY

Argonaute proteins repress gene expression and defend against foreign nucleic acids using short RNAs or DNAs to specify the correct target RNA or DNA sequence. We have developed single-molecule methods to analyze target binding and cleavage mediated by the Argonaute:guide complex, RISC. We find that both eukaryotic and prokaryotic Argonaute proteins reshape the fundamental properties of RNA:RNA, RNA:DNA, and DNA:DNA hybridization: a small RNA or DNA bound to Argonaute as a guide no longer follows the well-established rules by which oligonucleotides find, bind, and dissociate from complementary nucleic acid sequences. Moreover, mouse AGO2-RISC can distinguish between a miRNA-like target, to which it binds tightly, and the products of target cleavage, which it releases more quickly. By re-writing the rules for nucleic acid hybridization, Argonautes allow oligonucleotides to serve as specificity determinants with thermodynamic and kinetic properties more typical of RNA-binding proteins than of RNA or DNA.

INTRODUCTION

Small silencing RNAs such as microRNAs (miRNAs) and small interfering RNAs (siRNAs) direct Argonaute proteins to repress cellular mRNAs and silence foreign RNAs including viral mRNAs and RNA genomes. Argonaute proteins typically acquire their guide RNAs as 19–26 bp duplexes produced by dedicated processing and loading pathways (Hutvagner and Simard, 2008; Kawamata and Tomari, 2010). The final steps in production of a functional Argonaute-small RNA complex require expulsion of the passenger strand from the small RNA duplex to generate RISC (RNA-induced silencing complex)—an Argonaute protein bound to a single-stranded guide RNA (Matranga et al., 2005; Rand et al., 2005; Leuschner et al., 2006; Kwak and Tomari, 2012).

Argonaute proteins share a modular structure comprising the PAZ, Mid, and PIWI domains (Song et al., 2004). The Mid domain binds the 5′ phosphate of the guide RNA, anchoring it to the protein (Wang et al., 2008b). The Mid domain can also recognize the first base (g1) of a guide RNA. Mid domains that prefer a uridine at this position explain why miRNAs typically begin with U (Frank et al., 2010; Cora et al., 2014). One consequence of the Mid-domain:g1U interaction is that the g1 position of a miRNA, siRNA, or piRNA is unavailable for base pairing (Ma et al., 2005; Parker et al., 2005; Wang et al., 2014). The PAZ domain binds the 3′ end of the guide. PAZ domains play critical roles in loading Argonaute proteins with miRNA and siRNA duplexes during RISC assembly, and

have been proposed to compete for binding with target RNAs that can base-pair with the 3′ terminal nucleotides of the guide (Song et al., 2003; Lingel et al., 2004b; Ma et al., 2004; Tomari and Zamore, 2005). The PIWI domain, a structural homolog of the DNA-guided RNA endonuclease RNase H, contains the catalytic site, which positions a pair of magnesium ions near the scissile phosphate (Liu et al., 2004; Martinez and Tuschl, 2004; Schwarz et al., 2004; Wang et al., 2008b; Wang et al., 2009; Schirle et al., 2014). Some animal Argonaute proteins contain an additional N-terminal domain that prevents base pairing of the target to the guide beyond position guide position g16 (Kwak and Tomari, 2012; Faehnle et al., 2013; Hauptmann et al., 2013).

In animals, miRNAs and siRNAs typically silence gene expression by distinct mechanisms. Mammalian and fly miRNAs bind their targets via the seed sequence, a domain comprising guide nucleotides g2–g8 (Lewis et al., 2003; Rajewsky and Socci, 2004; Krek et al., 2005; Lewis et al., 2005; Lim et al., 2005). miRNA-directed Argonaute binding provides a platform that can recruit proteins that trigger exonucleolytic RNA degradation or inhibit translational initiation or elongation (Huntzinger and Izaurralde, 2011). Argonaute pre-positions these seven nucleotides so that they appear to be present in an RNA A-form helix, with its characteristically stacked bases, despite only one RNA strand being present in RISC (Ma et al., 2004; Parker et al., 2005; Wang et al., 2008a; Parker et al., 2009; Elkayam et al., 2012; Nakanishi et al., 2012; Wee et al.,

2012). By organizing the seed nucleotides into a conformation favorable for base-pairing (Ma et al., 2004; Parker et al., 2005; Wang et al., 2008a; Elkayam et al., 2012; Nakanishi et al., 2012), Argonaute pre-pays the entropic penalty inherent in binding complementary RNAs (Parker et al., 2009). Consequently, the binding specificity of RISC derives mainly from the seed sequence. In fact, mouse AGO2-RISC has nearly the same affinity for a seed-matched target as it does for a fully complementary target (Wee et al., 2012).

siRNAs generally have more extensive complementarity to their target RNAs than miRNAs. To act as endonucleases, Argonaute proteins must retain the key catalytic amino acids that organize the active site—a property of mammalian AGO2 but not AGO1 or AGO4 (AGO3 retains the catalytic residues but nonetheless cannot cleave RNA (Faehnle et al., 2013; Hauptmann et al., 2013)). Efficient endonucleolytic cleavage of a target requires that the guide base pair at least 5–8 nucleotides beyond the seed; this additional 'zippering' of the guide:target helix provides little additional binding energy, but allows the enzyme to attain a catalytically competent conformation (Haley and Zamore, 2004; Ameres et al., 2007) that places the scissile phosphate near a magnesium ion in the active site (Wang et al., 2008a; Wang et al., 2009) . Complementarity between g2–g16 and t2–t16 allows Argonaute to cleave the target phosphodiester bond between nucleotides t10 and t11 (Elbashir et al., 2001a; Elbashir et al., 2001b; Rivas et al., 2005).

In the absence of protein, the fundamental properties of ~21 nt RNA oligomers make them poor guides for directing gene regulation. Single-stranded, 5′ monophosphorylated RNA is readily degraded by endo- and exonucleases and can form intra- and intermolecular structures that inhibit target binding. At physiological temperature, pH, and ionic strength, 21 nt RNA oligomers bind with little specificity, because sub-sequences >12 nt hybridize stably to complementary sites in the transcriptome and because this high affinity binding accommodates insertions, deletions, and mismatches (Herschlag, 1991). Moreover, the rate of RNA and DNA hybridization is limited by the rate of successful collisions that convert to stable binding events (Ross and Sturtevant, 1960; Ross and Sturtevant, 1962; Nygaard and Hall, 1964; Wetmur and Davidson, 1968). This slow on-rate (k_{on}) means that the search for complementary targets is the rate-determining step for base pairing between 21 nt guides and their complementary targets. Once formed, 21 bp RNA:RNA duplexes are nearly irreversible in physiological conditions: a fully base-paired double-stranded RNA composed of let-7a and its complement is predicted to have a $K_D = 6.3 \times 10^{-7}$ nM, implying a $k_{off} = 5.7 \times 10^{-9}$ s^{-1} (τ = ~5.6 years). In contrast, an 8 bp duplex formed with just the let-7a seed sequence is unstable: the predicted K_D = 56 μM implies a $k_{off} = 52$ s^{-1} (τ = ~20 msec).

We sought to understand how Argonaute proteins overcome the inherent limitations posed by short oligonucleotide guides. Here, we show that

Argonaute proteins reshape the fundamental properties of RNA:RNA, RNA:DNA, and DNA:DNA hybridization. We find that once bound to Argonaute, a small RNA or DNA guide no longer follows the well-established rules for finding, binding, and dissociating from complementary nucleic acid sequences. By re-writing the rules, Argonautes allow oligonucleotides to serve as specificity determinants with thermodynamic and kinetic properties more typical of RNA-binding proteins than of nucleic acids.

RESULTS

Measurement of RNAi using single-molecule spectroscopy

To measure how Argonaute proteins alter the properties of oligonucleotides, we used Co-localization Single-Molecule Spectroscopy (CoSMoS), an implementation of multi-color, total internal reflection fluorescence (TIRF) microscopy that achieves high signal-to-noise ratios by exciting only those fluorescent molecules immediately above the slide surface (Friedman et al., 2006). To adapt CoSMoS to study RISC, a fluorescently labeled target RNA was attached to a glass surface via a biotin-streptavidin-biotin-PEG 3,400 linkage and then incubated with purified RISC assembled in vitro to contain a fluorescent guide strand (Figure 3.1A). The strategy relies on two novel reagents developed for these studies: (1) a target RNA designed to allow the unambiguous differentiation between target cleavage and photobleaching; and

(2) RISC assembled via the cellular Argonaute-loading pathway using an siRNA duplex containing a fluorescently labeled guide strand and then purified to remove unassembled siRNA and cleaved passenger strands (Flores-Jasso et al., 2013).

Figure 3.1

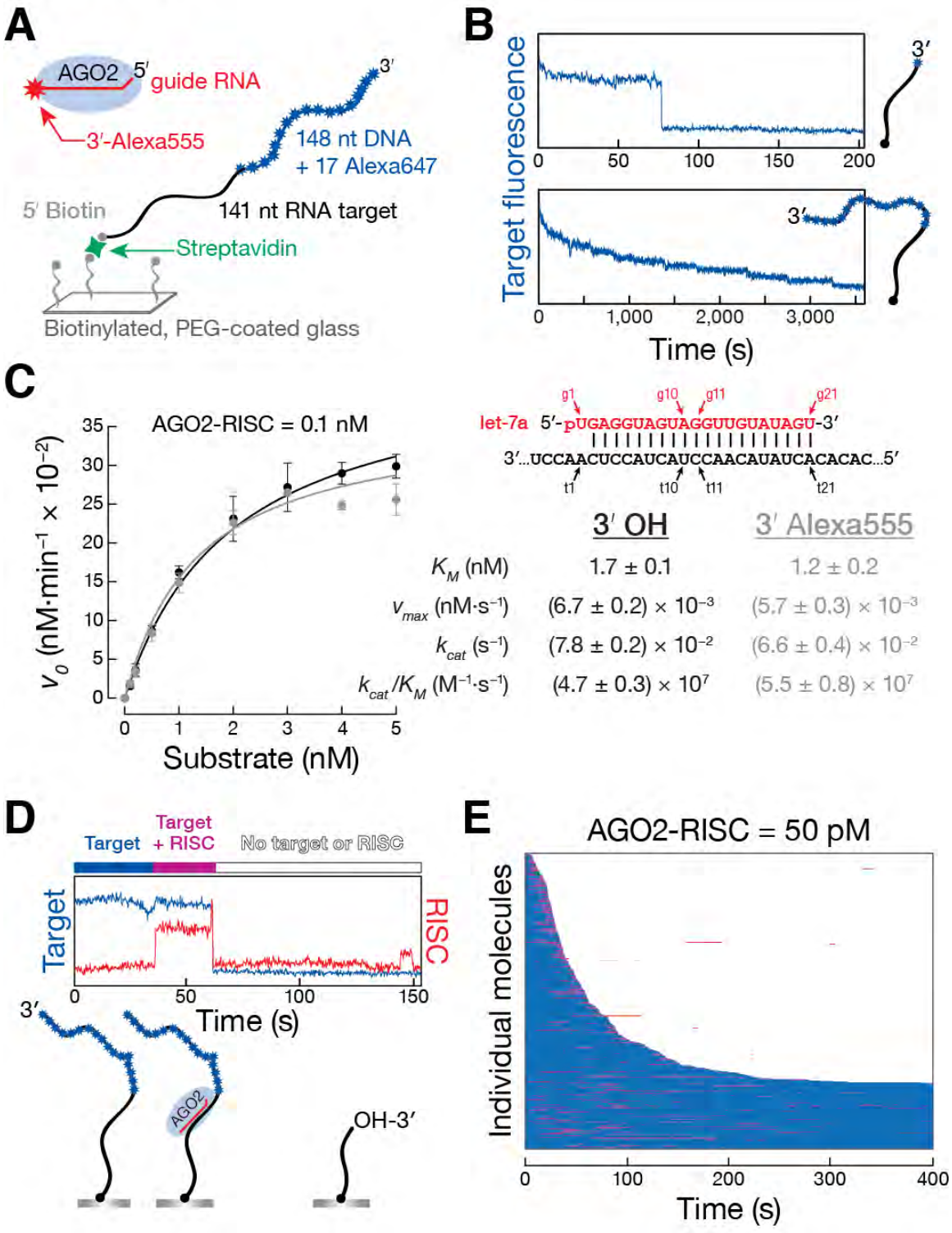

A
AGO2
5′
guide RNA
3′-Alexa555
3′
148 nt DNA
+ 17 Alexa647
5′ Biotin
141 nt RNA target
Streptavidin
Biotinylated, PEG-coated glass
B
Target fluorescence
3′
0
50
100
150
200
3′
0
1,000
2,000
3,000
Time (s)
C
AGO2-RISC = 0.1 nM
v_0 (nM·min^{-1} × 10^{-2})
35
30
25
20
15
10
5
0
0
1
2
3
4
5
Substrate (nM)
g1
g10
g11
g21
let-7a 5′-pUGAGGUAGUAGGUUGUAUAGU-3′
3′…UCCAACUCCAUCAUCCAACAUAUCACACAC…5′
t1
t10
t11
t21
3′ OH
3′ Alexa555
K_M (nM)
1.7 ± 0.1
1.2 ± 0.2
v_{max} (nM·s^{-1})
(6.7 ± 0.2) × 10^{-3}
(5.7 ± 0.3) × 10^{-3}
k_{cat} (s^{-1})
(7.8 ± 0.2) × 10^{-2}
(6.6 ± 0.4) × 10^{-2}
k_{cat}/K_M (M^{-1}·s^{-1})
(4.7 ± 0.3) × 10^7
(5.5 ± 0.8) × 10^7
D
Target
Target + RISC
No target or RISC
Target
RISC
0
50
100
150
Time (s)
3′
AGO2
OH-3′
E
AGO2-RISC = 50 pM
Individual molecules
0
100
200
300
400
Time (s)

Figure 3.1: Single-Molecule Analysis of Nucleic Acid-Guided Argonaute Proteins

(A) Strategy to measure RNA- or DNA-guided Argonaute interactions with RNA or DNA targets. (B) Photobleaching of a target labeled with a single Alexa647 dye is indistinguishable from target cleavage. In contrast, the stepwise photobleaching of a target with 17 Alexa647 dyes is readily distinguished from target cleavage. (C) Michaelis-Menten analysis of target cleavage for a standard RNA guide and a 3′ Alexa555-labeled RNA guide. Mean ± S.D. (n = 3). (D) A trace of an individual molecule of target RNA undergoing RNAi.
Blue: 5′ -tethered, 3′ (Alexa647 × 17)-labeled RNA target, fully complementary to let-7a; red: mouse AGO2-RISC programmed with 3′ Alexa555-labled let-7a RNA guide. Colored bars above trace summarize the species observed. This color code is used in the rastergrams. (E) Color-coded rastergram representation of let-7a-guided AGO2 binding and cleaving a fully complementary RNA target. The rastergram presents 426 individual RNA target molecules, each in a single row.

Photobleaching of fluorescent molecules is a technical challenge that plagues many single-molecule experiments, especially when high time resolution is required or when a molecule of interest must be continuously excited with laser light for an extended time. To overcome photobleaching and to distinguish photobleaching from target cleavage, we constructed a 141 nt RNA target containing 17 Alexa647 dyes within a 148 nt DNA 3′ extension. The multiply labeled target provided two related advantages. First, its extreme brightness allowed the use of decreased laser power, thereby decreasing the rate at which individual dyes photobleached. This allowed long observation times (30 min continuous illumination capturing 10,000 frames at 100 ms per frame; Figure 3.1B and Materials and Methods). Second, the presence of multiple Alexa647 dyes yielded a characteristic stepwise photobleaching pattern that was readily distinguishable from the all-or-none fluorescence change caused by target cleavage and 3′ product release. Figure 3.1B compares two molecules undergoing photobleaching: the target labeled with a single 3′ Alexa647 dye undergoes binary signal loss indistinguishable from target cleavage, whereas the target bearing 17 Alexa647 dyes gradually loses fluorescence in many discrete steps.

Mouse AGO2 was loaded with an RNA guide in cytoplasmic extract from *Ago2*$^{-/-}$ mouse embryonic fibroblasts overexpressing AGO2 under the control of the murine stem cell virus promoter (O'Carroll et al., 2007). Loading was

accomplished using a double-stranded siRNA carrying a 3′ Alexa555 dye on the guide strand; programmed RISC was then sequence-affinity purified (Flores-Jasso et al., 2013). To test whether dye addition altered the properties of AGO2-RISC, we compared the K_M and k_{cat} of AGO2 programmed with an unmodified guide corresponding to the sequence of let-7a, with AGO2 programmed with the 3′ Alexa555-labeled guide (Figures 3.1C and S3.1A). The K_M for let-7a-loaded RISC (1.7 ± 0.1 nM) was nearly identical to that containing 3′ Alexa555-labeled let-7a (1.2 ± 0.2 nM). Moreover, these K_M values agree well with previous values for human AGO2-RISC (Martinez and Tuschl, 2004; Rivas et al., 2005; Ameres et al., 2007). Similarly, k_{cat} for RISC containing the 3′ Alexa555-labeled let-7a guide ($6.6 \pm 0.4 \times 10^{-2}$ sec^{-1}) was similar to the k_{cat} of RISC programmed with let-7a without the dye ($7.8 \pm 0.2 \times 10^{-2}$ sec^{-1}).

RISC Changes the Rate-Determining Step for Nucleic Acid Hybridization

Argonaute proteins have been proposed to increase the rate of nucleic acid hybridization by pre-organizing the nucleotides of the seed sequence into a stacked conformation that makes productive collisions with target more likely. The association rate constant, k_{on}, for mammalian AGO2 has been inferred from K_D and k_{off} values measured in ensemble binding experiments (Wee et al., 2012) or estimated by fitting pre-steady state ensemble data to a three-phase exponential model in which the fastest phase was assumed to correspond to k_{on} (Deerberg et al., 2013).

To measure k_{on} directly, we simultaneously recorded the fluorescence of individual target RNAs attached to the slide and individual molecules of mouse AGO2-RISC containing fluorescent guide strand (Figure 3.1D and Materials and Methods). For each target RNA molecule, RISC arrival time was taken to be the first detectable co-localization of RISC fluorescence and target RNA fluorescence. Our standard conditions detected the arrival of RISC molecules that remained co-localized with a target ≥200 msec (one frame at 5 frames·s^{-1}; Materials and Methods). Figure 3.1D provides an example of Alexa555-labeled RISC arriving at an Alexa647-labeled target: when RISC arrives at ~40 s, the Alexa 555 fluorescence co-localizing with the Alexa647 target increases in a single step; it remains high until both Alexa555 (RISC) and Alexa647 (target) fluorescence drop to baseline at ~60 s, signifying target cleavage and simultaneous departure of RISC and the 3′ cleavage product. Figure 3.1E displays 426 individual single-molecule traces, ordered by time of target cleavage, as a 'rastergram.' Rastergrams summarize the arrivals, departures, and target cleavage events for many individual target molecules.

To understand how AGO2 changes the rate at which an oligonucleotide arrives at a target, we used CoSMoS to compare the k_{on} of single-stranded let-7a RNA and let-7a-programmed AGO2-RISC (Figure 3.2). After their arrival at the target, let-7a alone and let-7a bound to AGO2 follow different paths. Formation of a 21 bp RNA:RNA duplex is essentially irreversible under

physiological conditions, so observation of let-7a ended when its Alexa555 label photobleached. Subsequently, no new, fluorescent RISC was detected binding to the target. The target, labeled with 17 Alexa647 dyes, continued to be detectable, gradually losing fluorescence via discrete photobleaching events (Figure 3.2A). In contrast, binding of let-7a RISC ended with target cleavage; Alexa555 and Alexa647 fluorescence were lost simultaneously.

On-rates for RISC (k_{on}) or let-7a alone were determined by fitting the cumulative distribution of arrivals to a single exponential, corrected for non-specific background binding to the slide (Figure S3.1B and Materials and Methods). The on-rate of let-7a RNA alone binding to a fully complementary target ($9.1 \pm 1.7 \times 10^6$ $M^{-1}{\cdot}s^{-1}$; Figures 3.2A and S3.2) was considerably slower than the rate of macromolecular diffusion. The sequence of let-7a comprises only three (A, G, and U) of the four nucleotides. The on-rate for let-7a alone measured by CoSMoS agrees well with previous single-molecule estimates of k_{on} for short oligonucleotides lacking G or C (Zhang et al., 2014). In comparison, k_{on} for let-7a-programmed mouse AGO2-RISC binding to the fully complementary target RNA ($3.9 \pm 0.5 \times 10^8$ $M^{-1}{\cdot}s^{-1}$) was ~43-fold faster than k_{on} for let-7a alone (Figures 3.2A and S3.2–S3.3). Moreover, the rate at which AGO2-RISC finds its target approaches the limit of macromolecular diffusion at 37°C: ~6.4×10^9 $M^{-1}{\cdot}s^{-1}$ under our standard conditions, which include 20% glycerol (Segur and Oberstar, 1951; Berg and von Hippel, 1985). Thus, as

proposed previously (Wee et al., 2012; Deerberg et al., 2013; Jung et al., 2013), Argonaute accelerates productive arrival of its guide at a complementary target sequence.

While these data indicate that AGO2 improves k_{on} by ~43-fold for binding of let-7a to a fully complementary target sequence, the unusual sequence composition of let-7a might understate the general enhancement in target finding. To test this, we determined k_{on} for a miRNA containing all four nucleotides, miR-21, either alone or bound to mouse AGO2 (Figures 3.2 and S3.2–S3.3). As expected from its greater sequence complexity, miR-21 RNA alone bound its complementary RNA target ~17 times more slowly (k_{on}, miR-21 alone = $5.3 \pm 0.2 \times 10^5$ $M^{-1} \cdot s^{-1}$) than let-7a alone. Mouse AGO2 accelerated miR-21 binding ~250-fold (k_{on} miR-21 AGO2-RISC = $1.3 \pm 0.1 \times 10^8$ $M^{-1} \cdot s^{-1}$). Thus, accelerating oligonucleotide target finding to close to the rate of diffusion is a general property of AGO2-RISC.

Is this acceleration of target finding a general property of other Argonaute proteins or unique to mouse AGO2? To answer this, we measured k_{on} to a fully complementary sequence for TtAgo, the DNA-guided Argonaute protein (Swarts et al., 2014a) from the eubacterium *Thermus thermophilus*. *T. thermophilus* grows at 62–75°C (Cava et al., 2009), and TtAgo does not efficiently cleave either RNA or DNA at 37°C (Figure S3.4A). Control experiments established that the addition of an Alexa555 dye to the 3′ end of the DNA guide does not alter

the ensemble binding properties of TtAgo (Figure S3.4B). In vivo, TtAgo binds 16 nt DNA guides (SMJ and PDZ, unpublished), so we loaded TtAgo at 75°C with a single-stranded DNA comprising the first 16 nt of let-7a, and then studied its binding at 37°C using CoSMoS. On its own, the 16 nt "let-7a" guide bound a complementary RNA target ~140 times more slowly (k_{on} = 4.6 ± 0.1 × 10^5 $M^{-1}{\cdot}s^{-1}$) than the same DNA guide bound to TtAgo (k_{on} = 6.2 ± 0.1 × 10^7 $M^{-1}{\cdot}s^{-1}$; Figures 3.2 and S3.4C).

Why is the k_{on} for *Thermus thermophilus* Argonaute ~6 times slower than that of mouse AGO2? We suspect that at 37°C TtAgo spends less time in a binding-competent conformation, consistent with its greater target cleavage activity at 75°C (Figure S3.4A). We conclude that both mouse AGO2 and TtAgo alter the rate-determining step for nucleic acid hybridization (k_{on}), ensuring that the speed at which Argonaute finds its complementary target RNA or DNA is limited by the rate of macromolecular diffusion.

Figure 3.2

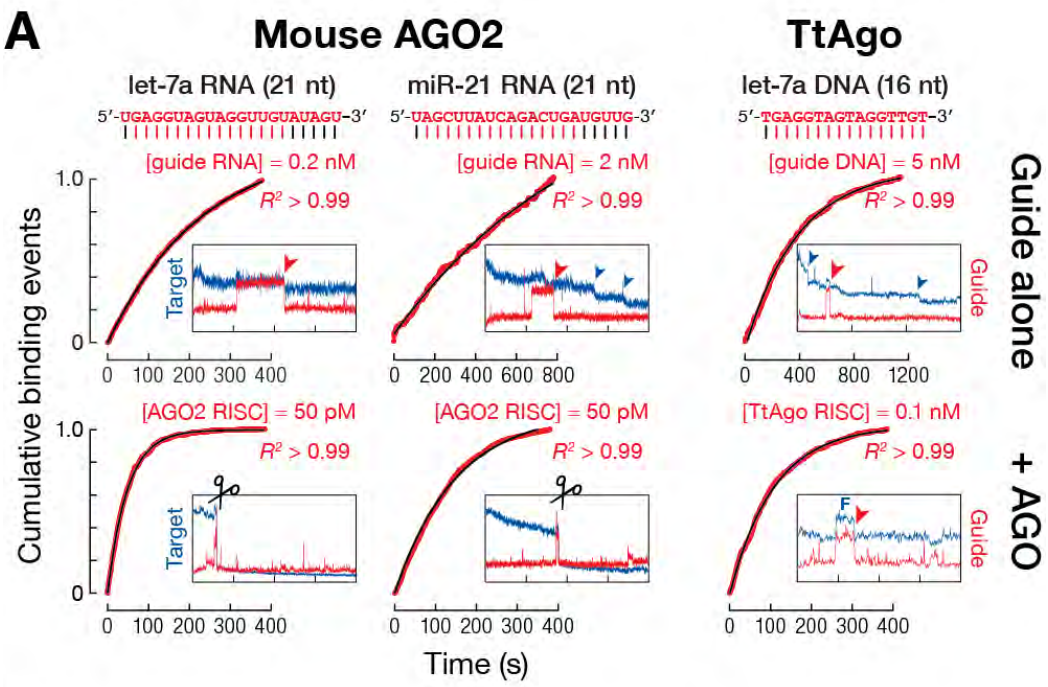

A
Mouse AGO2
TtAgo
let-7a RNA (21 nt)
miR-21 RNA (21 nt)
let-7a DNA (16 nt)
5′-UGAGGUAGUAGGUUGUAUAGU-3′
5′-UAGCUUAUCAGACUGAUGUUG-3′
5′-TGAGGTAGTAGGTTGT-3′
[guide RNA] = 0.2 nM
[guide RNA] = 2 nM
[guide DNA] = 5 nM
[AGO2 RISC] = 50 pM
[AGO2 RISC] = 50 pM
[TtAgo RISC] = 0.1 nM
R² > 0.99
Target
Guide
Guide alone
+ AGO
Cumulative binding events
Time (s)
B
Seed Sequence
Guide
Target
kon (M⁻¹·s⁻¹)
Guide alone
+ AGO
mouse AGO2
TtAgo
9.1 ± 1.7 × 10⁶
3.9 ± 0.5 × 10⁸
4.6 ± 1.6 × 10⁶
2.4 ± 0.1 × 10⁸
5.1 ± 1.8 × 10⁶
2.8 ± 0.5 × 10⁸
7.6 ± 0.9 × 10⁶
3.6 ± 0.2 × 10⁷
2.1 ± 0.1 × 10⁶
1.0 ± 0.1 × 10⁹
7.8 ± 0.2 × 10⁸
6.4 ± 0.2 × 10⁸
5.3 ± 0.2 × 10⁵
1.3 ± 0.1 × 10⁸
4.7 ± 0.1 × 10⁵
6.4 ± 0.1 × 10⁷
7.1 ± 0.1 × 10⁷
4.8 ± 0.1 × 10⁷
4.6 ± 0.1 × 10⁵
6.2 ± 0.1 × 10⁷
2.0 ± 0.1 × 10⁷
5.8 ± 0.1 × 10⁶

Figure 3.2: Compared to Nucleic Acid Alone, Argonaute Accelerates Guide Binding to Target

(A) Comparison of target binding rates (kon) by 21 nt RNA-guided mouse AGO2- and 16 nt DNA-guided TtAgo versus the RNA or DNA guide strands alone. Cumulative binding fraction plots are accompanied by the fluorescence intensity trace for a representative individual molecule. Red arrowheads: photobleaching of the Alexa555 guide; blue arrowheads: stepwise photobleaching of a single Alexa647 group; scissors: Argonaute-catalyzed target cleavage; blue F: Förster resonance energy transfer from the Alexa555 guide to the Alexa647 target bearing 17 dye moieties. (B) Comparison of kon values for mouse AGO2-let-7a and TtAgo-let-7a. For let-7a and miR-21 RNA targets, mouse AGO2 and RNA alone values are reported as mean ± S.D. (n = 3), with >1,000 individual molecules collected. All other values were measured using several hundred individual molecules, and error of fit is reported. k_{on} values were corrected for the rate of non-specific binding to the slide surface.

Argonaute Accelerates the Rate of Target Finding by Creating the Seed Sequence

The three structural domains of Argonaute proteins divide their guide RNAs into discrete functional domains. To determine which domain contributes most to the Argonaute-dependent enhancement of target binding, we measured k_{on} using three different target RNAs: (1) a target complementary just to the seed sequence (g2–g8); (2) a target complementary to both the seed and the region of 3′ supplementary pairing (g13–g16); and (3) a target with complete complementarity to the guide (g2–g21; Figures 3.2B and S3.2–S3.3). For each target RNA, we determined k_{on} for both the guide alone and the guide loaded into mouse AGO2 (Figures 3.2 and S3.2–S3.3). We also measured k_{on} for let-7a-RISC binding to an RNA having ≤ 6 nt complementary to any region of the let-7a guide sequence and ≤ 4 nt complementary to the let-7a seed sequence. For this essentially non-complementary control RNA, we were unable to detect any binding interactions above non-specific background binding to the slide (Figure S3.5A).

Structural comparisons of eubacterial and human AGO2 show that an N-terminal Argonaute domain prevents pairing beyond g16 in animal Argonautes; and computational analyses of piRNAs in flies, silk moths, and mice suggest that target cleavage does not require complementarity beyond target position t16 (Wang et al., 2009; Kwak and Tomari, 2012; Wee et al., 2012;

Faehnle et al., 2013; Hauptmann et al., 2013; Wang et al., 2014). Thus even targets with complete complementarity to the guide strand are unlikely to pair past g16 when the guide is bound to Argonaute.

In the absence of protein, nucleic acid hybridization is favored by greater complementarity, presumably because the larger number of potential base pairs provides more opportunities for nucleation, the rate-determining step for productive binding (Egli and Saenger, 1988). Consistent with this principle, k_{on} for let-7a RNA alone increased approximately 2-fold, from $4.6 \pm 1.6 \times 10^6$ M^{-1} s^{-1} for the target with complementarity only to the seed sequence (g1–g8) or $5.1 \pm 1.8 \times 10^6$ M^{-1} s^{-1} for the target with seed and 3′ supplementary pairing (g1–g8 plus g13–g16) to $9.1 \pm 0.1 \times 10^6$ M^{-1} s^{-1} for the fully complementary target (g1–g21) (Figures 3.2B and S3.2). Yet when loaded in AGO2-RISC, let-7a bound all three targets with very similar, near diffusion-limited on-rates, ranging from $2.4 \pm 0.1 \times 10^8$ $M^{-1}{\cdot}s^{-1}$ to $3.9 \pm 0.5 \times 10^8$ $M^{-1}{\cdot}s^{-1}$; (Figures 3.2B and S3.2). Additionally, there was little to no difference in k_{on} when the total length of the target changed (Figure S3.6); let-7a AGO2-RISC had similar k_{on} values when comparing a the longer 141 nt RNA + 148 nt DNA target to 28 nt RNA targets that perfectly paired ($2.3 \pm 0.1 \times 10^8$ $M^{-1}{\cdot}s^{-1}$) or paired only to the seed nucleotides ($1.5 \pm 0.1 \times 10^8$ $M^{-1}{\cdot}s^{-1}$). In contrast, the apparent rate of RISC finding an RNA target fully complementary to let-7a except for the seed nucleotides ($3.6 \pm 0.2 \times 10^7$ $M^{-1}{\cdot}s^{-1}$) was ~10-fold slower (Figures 3.2B and

S3.5B). We observed similar k_{on} effects for miR-21 RISC, whose k_{on} was ~20-fold faster than miR-21 RNA alone when binding a target with complementarity only to the seed ($1.1 \pm 0.1 \times 10^7$ $M^{-1} \cdot s^{-1}$) or to both the seed and four, 3′ supplementary bases ($1.7 \pm 0.1 \times 10^7$ $M^{-1} \cdot s^{-1}$; Figure S3.2). We conclude that the seed sequence created by mouse AGO2 accounts for most of the enhancement in the rate of target finding.

To further test this idea, we measured k_{on} for a series of six target RNAs bearing a dinucleotide mismatch in their seed-complementary sequence (Figures 3.3A, S3.3 and S3.5B). We performed these experiments at 10 frames·s^{-1}, a time resolution likely to be sufficient to detect the first arrival of RISC, because k_{on} values did not change at higher time resolutions (25 frames·s^{-1}; data not shown). Compared to a seed-matched target, dinucleotide mismatches at guide positions g2g3, g3g4, g4g5, or g5g6 reduced k_{on} 6.3- to 10-fold. Mismatches with positions g6g7 or g7g8 reduced k_{on} just 1.3-fold, compared to a target complementary to the 7 nt seed sequence. These data further support the view that Argonautes accelerate target finding by pre-organizing the seed and that acceleration is diminished when the seed pairing is disrupted at positions g2–g5.

TtAgo also required seed complementarity to accelerate the rate of target finding: the rate for TtAgo, guided by a 16 nt DNA, to find a target DNA complementary to the seed ($k_{on} = 7.1 \pm 0.1 \times 10^7$ $M^{-1} \cdot s^{-1}$) or both the seed and

four, 3′ supplementary nucleotides (k_{on} = 4.8 ± 0.1 × 10^7 $M^{-1}{\cdot}s^{-1}$), was essentially the same as when the entire guide was complementary to the target (k_{on} = 6.4 ± 0.1 × 10^7 $M^{-1}{\cdot}s^{-1}$; Figures 3.2 and S3.2, Table 3.1).

Figure 3.3

A

Seed Sequence	k_{on} ($M^{-1}s^{-1}$) Guide alone	k_{on} ($M^{-1}s^{-1}$) + AGO	k_{on} relative to seed
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|||||||UAGGUUGUAUAGU	$4.6 \pm 1.6 \times 10^{6}$	$2.4 \pm 0.1 \times 10^{8}$	1
5′-UGAGGUAGUAGGUUGUAUAGU-3′ AGA|||||UAGGUUGUAUAGU		$3.8 \pm 0.1 \times 10^{7}$	0.16 ± 0.01
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|AG||||UAGGUUGUAUAGU		$3.4 \pm 0.2 \times 10^{7}$	0.14 ± 0.01
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A||AA|||UAGGUUGUAUAGU		$2.4 \pm 0.1 \times 10^{7}$	0.10 ± 0.01
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|||GU||UAGGUUGUAUAGU		$3.9 \pm 0.1 \times 10^{7}$	0.16 ± 0.01
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A||||UA|UAGGUUGUAUAGU		$2.0 \pm 0.1 \times 10^{8}$	0.81 ± 0.03
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|||||AGUAGGUUGUAUAGU		$2.0 \pm 0.1 \times 10^{8}$	0.82 ± 0.03
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A||||||GUAGGUUGUAUAGU		$1.8 \pm 0.1 \times 10^{8}$	0.74 ± 0.02

B

Seed Sequence	k_{off} (s^{-1}) Guide alone (predicted)	k_{off} (s^{-1}) + AGO (measured)	k_{off} relative to seed	K_D (nM)
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|||||||||||||||||||||	5.7×10^{-9}	N.D.	N.D.	N.D.
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|||||||UAGG||||AUAGU	0.72	0.0030 ± 0.0004	1.0 ± 0.1	0.011 ± 0.002
5′-UGAGGUAGUAGGUUGUAUAGU-3′ AAAAAAAAAA|||||||||||	0.60	0.79 ± 0.08	220 ± 30	22 ± 3
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|||||||UAGGUUGUAUAGU	52	0.0036 ± 0.0003	1	0.015 ± 0.002
5′-UGAGGUAGUAGGUUGUAUAGU-3′ AGA|||||UAGGUUGUAUAGU		11 ± 1	3,000 ± 300	280 ± 40
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|AG||||UAGGUUGUAUAGU		7.3 ± 0.5	2,000 ± 200	220 ± 30
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A||AA|||UAGGUUGUAUAGU		6.0 ± 0.4	1,700 ± 200	250 ± 40
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|||GU||UAGGUUGUAUAGU		12 ± 1	3,200 ± 300	300 ± 40
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A||||UA|UAGGUUGUAUAGU		3.5 ± 0.1	960 ± 90	18 ± 2
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A|||||AGUAGGUUGUAUAGU		0.24 ± 0.01	67 ± 6	1.2 ± 0.2
5′-UGAGGUAGUAGGUUGUAUAGU-3′ A||||||GUAGGUUGUAUAGU		0.086 ± 0.002	24 ± 2	0.48 ± 0.06

Figure 3.3: Mismatches Highlight the Role of the Seed Sequence in Target Binding

Comparison of the kinetic properties of let-7a-guided mouse AGO2-RISC with different targets. Values were derived from data collected from several hundred individual RNA target molecules; error of fit is reported. All k_{on} and k_{off} values were corrected for the rate of non-specific binding to the slide surface. k_{off} and K_D values for nucleic acid in the absence of protein were predicted from the measured k_{on} and $\Delta G_{37°C}$ of binding calculated by nearest neighbor analysis (Reuter and Mathews, 2010; Turner and Mathews, 2010). k_{off} was not determined for the fully complementary target because it was cleaved under our experimental conditions.

Table 3.1

21 nt let-7a RNA loaded in mouse AGO2			
	RNA Target		
Extent of complementarity	**k_{on} ($M^{-1}·s^{-1}$)**	**k_{off} (s^{-1})**	**K_D (nM)**
Complete	$3.9 \pm 0.5 \times 10^8$	<< PB	<< PB
Seed only	$2.4 \pm 0.1 \times 10^8$	0.0036 ± 0.0003	0.015 ± 0.002
Seed + 3′ Supplementary	$2.8 \pm 0.5 \times 10^8$	0.0030 ± 0.0004	0.011 ± 0.002
	DNA Target		
Complete	$1.0 \pm 0.1 \times 10^9$	<< PB	<< PB
Seed only	$7.8 \pm 0.2 \times 10^8$	0.41 ± 0.09	0.53 ± 0.07
Seed + 3′ Supplementary	$6.4 \pm 0.2 \times 10^8$	0.57 ± 0.02	0.9 ± 0.1

16 nt let-7a-derived DNA guide loaded in TtAgo			
	RNA Target		
Extent of complementarity	**k_{on} ($M^{-1}·s^{-1}$)**	**k_{off} (s^{-1})**	**K_D (nM)**
Complete	$6.2 \pm 0.1 \times 10^7$	<< PB	<< PB
Seed only	$2.0 \pm 0.1 \times 10^7$	0.35 ± 0.03	17 ± 2
Seed + 3′ Supplementary	$5.8 \pm 0.1 \times 10^6$	0.51 ± 0.06	88 ± 12
	DNA Target		
Complete	$6.4 \pm 0.1 \times 10^7$	<< PB	<< PB
Seed only	$7.1 \pm 0.1 \times 10^7$	1.0 ± 0.1	15 ± 2
Seed + 3′ Supplementary	$4.8 \pm 0.1 \times 10^7$	0.86 ± 0.08	18 ± 2

Table 3.1: Kinetic values for RISC with RNA or DNA Targets

Properties of Argonaute-guide complexes binding to RNA or DNA targets with different extents of complementarity to the guide strand. << PB, not determined because koff was much slower than the rate of photobleaching.

The rate of a diffusion-limited bimolecular reaction depends on the viscosity of the solution, because diffusion proceeds more slowly at higher viscosity (Berg and von Hippel, 1985). To test whether target finding by AGO2 similarly depends on viscosity, we measured k_{on} for let-7a-RISC without or with 20%—our standard conditions—or with 50% (w/v) glycerol (Figure 3.5). As predicted for a diffusion-limited process, the k_{on} for a fully complementary let-7a target increased from $3.9 \pm 0.5 \times 10^8$ $M^{-1} \cdot s^{-1}$ in 20% glycerol to $6.4 \pm 0.1 \times 10^8$ $M^{-1} \cdot s^{-1}$ when glycerol was omitted. When the glycerol concentration was increased to 50%, increasing the viscosity of the reaction, k_{on} decreased to $1.7 \pm 0.1 \times 10^8$ $M^{-1} \cdot s^{-1}$.

Figure 3.4

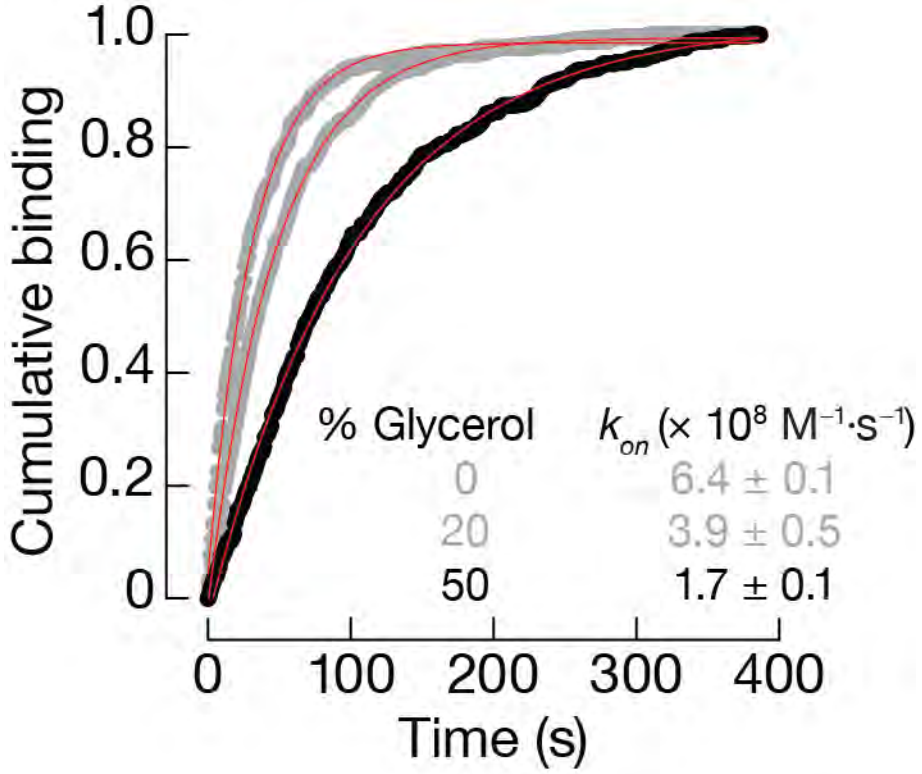

Cumulative binding
1.0
0.8
0.6
0.4
0.2
0
% Glycerol
k_{on} (× 10^8 $M^{-1}·s^{-1}$)
0
6.4 ± 0.1
20
3.9 ± 0.5
50
1.7 ± 0.1
0
100
200
300
400
Time (s)

Figure 3.4: RISC target finding is a diffusion-limited reaction

Cumulative binding curves for mouse let-7a AGO2-RISC to perfect targets in buffer containing 0, 20, or 50% (f.c.) glycerol. k_{on} represents the analysis of 521 targets for 0% glycerol, 1,264 targets for 20% glycerol, and 491 targets for 50% glycerol.

Seed Mismatches Cause Rapid Dissociation of Mouse AGO2 RISC

Ensemble experiments at 25°C show that mouse AGO2-RISC departs slowly from seed-matched targets (Wee et al., 2012), a time-scale too long for direct observation by fluorescence of individual RISC molecules, because photobleaching of the guide RNA dye generally occurs before a departure is observed (the rate of Alexa555 photobleaching under our standard condition was ~0.06 s^{-1}, τ ~ 17 s; Materials and Methods). As an alternative strategy to measure k_{off} for mouse AGO2-RISC at 37°C, a more physiologically appropriate temperature, we measured the apparent k_{off} over a range of laser exposure (i.e., by changing the frame length) and extrapolated to no laser exposure (the *y*-intercept) to obtain k_{off}: 0.0036 ± 0.0003 s^{-1}, τ ~280 s (Figure S3.5C). Because the photobleaching rate was much slower than the dissociation rate for less complementary targets, k_{off} was readily measured by standard methods for the six targets containing a dinucleotide mismatch within the let-7a seed-match (Figures 3.3B, S3.3, and S3.7). Compared to the seed-matched target, RISC dissociated from these seed-mismatched targets from 70 to 3,200 times faster. As we observed for k_{on}, individual positions within the seed contributed differentially to anchoring AGO2-RISC on the target RNA, with base pairs at g2–g6 contributing more than base pairs at g7 or g8. Even a single nucleotide mismatch to g8, converting a 7 nt seed to a 6 nt seed, increased k_{off} by 24-fold (Figures 3.3B and S3.3). Thus, RISC discriminates between seed-

matched and seed-mismatched targets both during its initial search and after it has bound; it finds seed-mismatched targets more slowly and remains bound to them for less time than fully seed-complementary targets.

Seed Pairing Explains Mouse AGO2 Binding to miRNA-Like Targets

The effect of target:guide mismatches within the seed sequence cannot be accurately predicted from nearest-neighbor thermodynamic rules (Xia et al., 1998). For example, a dinucleotide mismatch at the end of a 7 nt seed-match is predicted to increase ΔG for RISC:target binding by 1.7 (g2g3:t2t3 mismatch) to 1.9 kcal·mol^{-1} (g7g8 mismatch), compared to a fully complementary seed match target site (g2–g8; Figure S3.7). In reality the effect is much larger: a g2g3 dinucleotide mismatch (i.e., only g4–g8 paired) reduced the stability of the RISC:target complex by 6.1 ± 0.1 kcal·mol^{-1}, while a g7g8 mismatch (g2–g6 paired) reduced the stability of RISC binding by 2.7 ± 0.1 kcal·mol^{-1} (Figure S3.7). Similarly, a g3g4 mismatch, which is predicted to decrease binding stability by 3.8 kcal·mol^{-1}, reduced it by 5.9 ± 0.1 kcal·mol^{-1}.

Additional base pairs with the 3′ half of the guide (3′ supplementary pairing) are associated with high probability miRNA-binding sites (Grimson et al., 2007). The addition of four 3′ supplementary base pairs is predicted to change the $\Delta G_{37°C}$ of a fully base paired let-7a seed by –3.7 kcal·mol^{-1}, yet our experiments failed to detect a substantial change in binding stability for let-7a bound to its target by seven seed base pairs or by seven seed plus four

additional 3′ supplemental base pairs ($\Delta\Delta G_{37°C}$ = –0.20 ± 0.2 kcal·mol^{-1}). In contrast, dissociation of AGO2 guided by miR-21, which has a more AU-rich seed than let-7, was slowed >7-fold by adding 3′ supplementary base pairing (Figure S3.5D). Increasing the base-pairing strength of the seed by replacing three seed-match adenosines with 2,6 diaminopurine nucleotides decreased k_{off} 4-fold and increased RISC affinity by –0.83 kcal·mol^{-1}, far less than the predicted –3 kcal·mol^{-1} (Figure S3.5D ;Gryaznov and Schultz, 1994; Freier and Altmann, 1997).

AGO2 Discriminates between RNA and DNA targets

Mouse AGO2, like all known animal Argonautes, has only been reported to function by binding RNA targets. In contrast, TtAgo can cleave both RNA and DNA targets, although only DNA targets have been identified in vivo (Wang et al., 2008a; Wang et al., 2008b; Wang et al., 2009; Swarts et al., 2014a). How do animal Argonaute proteins discriminate between RNA and DNA? We compared the binding of mouse AGO2 to RNA targets with binding to the same sequences composed of DNA (Figures 3.2B and S3.2). As we observed for RNA, AGO2 accelerated the search for a complementary binding site on DNA. In fact, k_{on} for a seed-matched (g2–g8), seed-matched plus supplementary pairing (g2–g8, g13–g16), or a completely complementary (g2–g16) DNA target was ~2.3 to 3.3 times faster than the on-rate for the corresponding RNA target (Table 3.1). However, mouse AGO2-RISC did not remain stably bound to the DNA,

dissociating, on average, just ~2.4 s (k_{off} = 0.41 ± 0.09 s^{-1}) after binding a seed-matched DNA target. In contrast, AGO2-RISC remained bound to an otherwise identical RNA target for an average of ~280 s (k_{off} = 0.0036 ± 0.0003 s^{-1}; Table 3.1 and Figures S3.3 and S3.5C). We observed a similar difference in k_{off} for DNA and RNA targets with both seed and 3′ supplementary complementarity. The >110-fold faster dissociation of AGO2-RISC from DNA compared to RNA supports the view that even when acting in the nucleus, eukaryotic RISCs bind nascent transcripts, not single-stranded DNA (Buhler et al., 2006; Sabin et al., 2013).

In contrast to mammalian Argonautes, bacterial Argonautes are thought to preferentially bind and cleave foreign DNA, such as horizontally transferred plasmids (Olovnikov et al., 2013; Swarts et al., 2014a). Consistent with this function, TtAgo showed no substantive preference for binding RNA over DNA targets (Figures 3.2B, S3.2, and Table 3.1). TtAgo found its binding sites in RNA and DNA at similar rates (e.g., seed-matched target k_{on} = 2.0 ± 0.1 × 10^7 $M^{-1}{\cdot}s^{-1}$ for RNA versus 7.1 ± 0.1 × 10^7 $M^{-1}{\cdot}s^{-1}$ for DNA; Figures 3.2B and S3.2), and, once bound, departed from RNA and DNA at similar rates (k_{off} = 0.35 ± 0.03 s^{-1} for RNA versus 1.0 ± 0.1 s^{-1} for DNA; Figure S3.4C and Table 3.1).

A Kinetic Framework for Mammalian RNAi

For a 5′-tethered target, target cleavage by RISC leaves the 5′ cleavage product tethered to the slide surface, allowing detection of RISC that remains bound via

guide nucleotides g11–g21. When let-7a guided mouse AGO2, target cleavage and release of the 5′ cleavage product from RISC were simultaneous within the time resolution of our experiments (e.g., Figure 3.1D). This suggests that 5′ product release is faster than RISC dissociation from the 3′ cleavage product, which contains the seed-complementary sequence.

To measure 3′ cleavage product release, we synthesized a let-7a-complementary target with biotin on its 3′ end and 17 Alexa647 dyes at its 5′ end (Materials and Methods and Table S3.1). The 3′-tethered target allowed us to detect four distinct reaction species: (1) target alone, (2) RISC bound to the target, (3) RISC bound to the 3′ cleavage product, and (4) the 3′ product after RISC dissociation (Figure 3.5A). Our experiments with 5′-tethered target add information about RISC bound to 5′ cleavage product and the 5′ product alone, completing the set of all observable species in the RNAi reaction. Figure 3.5B presents rastergrams that summarize the results for hundreds of RNA target molecules where the reaction states of AGO2-RISC and target were observed for the entire duration of the experiment.

As expected, k_{on} for 3′- and 5′-tethered targets ($3.7 \pm 0.1 \times 10^8$ $M^{-1}{\cdot}s^{-1}$ and $3.9 \pm 0.5 \times 10^8$ $M^{-1}{\cdot}s^{-1}$, respectively) was nearly identical to let-7a-AGO2-RISC binding rates (Figures 3.2B and S3.2). However, the order and rates of dissociation differed considerably for the 5′ and 3′ cleavage products (Figure 3.5B). The first product to be released was nearly always the 5′ product,

after which AGO2-RISC slowly dissociated from the 3′ product. After AGO2-RISC departed, we frequently observed RISC rebinding to the 3′ cleavage product (Figure 3.5B). The 3′ product is complementary to the seed sequence; thus, let-7a AGO2-RISC maintains high affinity for a seed-match even after target cleavage, highlighting the essential role of the seed in RISC binding.

Figure 3.5

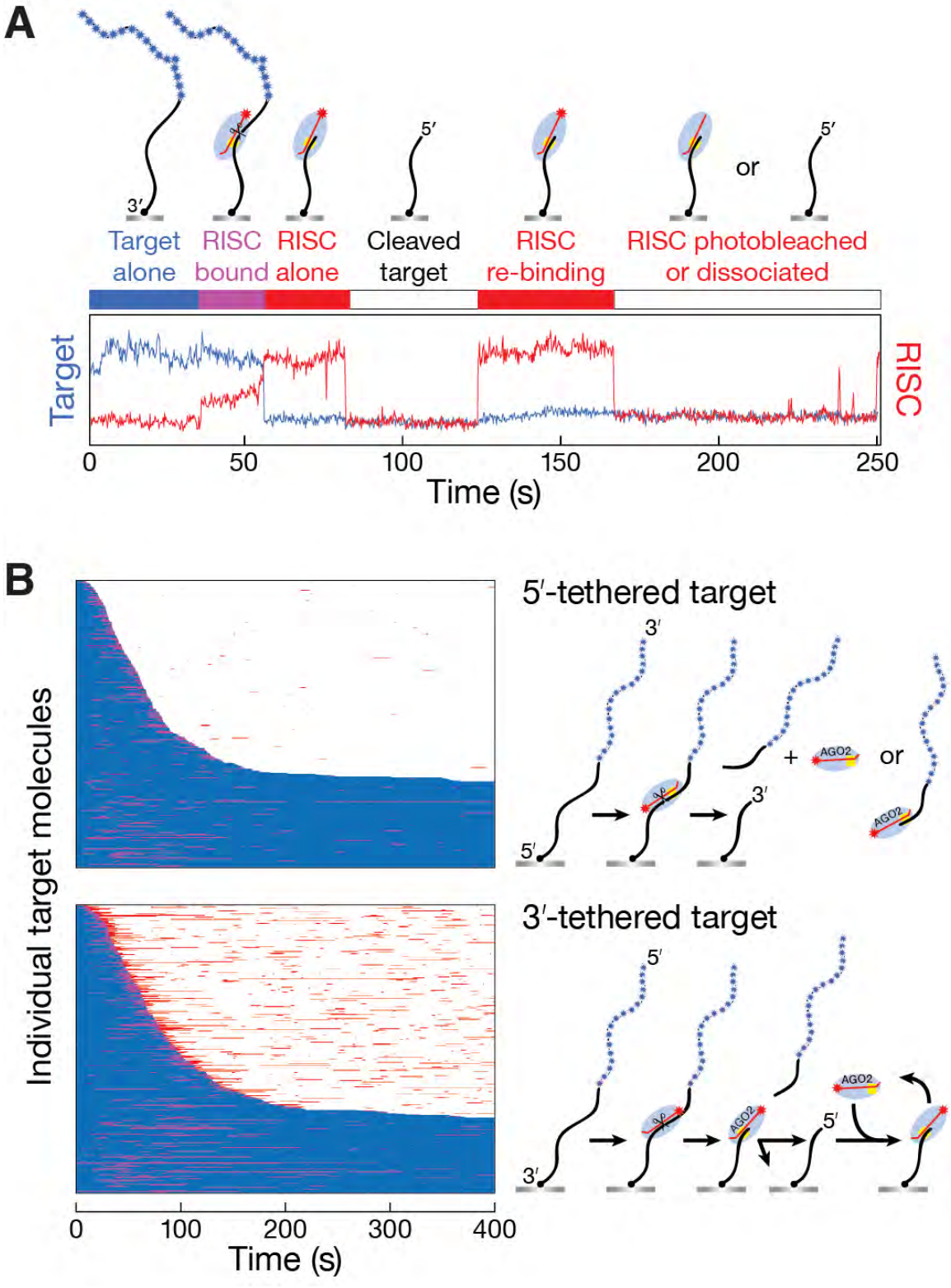

A
3′
5′
or
5′
Target alone
RISC bound
RISC alone
Cleaved target
RISC re-binding
RISC photobleached or dissociated
Target
RISC
0
50
100
150
200
250
Time (s)
B
Individual target molecules
5′-tethered target
3′
5′
3′
+
AGO2
or
AGO2
3′-tethered target
5′
3′
AGO2
AGO2
5′
0
100
200
300
400
Time (s)

Figure 3.5: let-7a Binds Tightly to the Seed-Matching, 3′ Product of Target Cleavage

(A) A trace of an individual molecule of a 3′-tethered target fully complementary to let-7a. The trace shows that mouse let-7a-RISC bound the target (magenta bar), cleaved the target, and then remained bound to the 3′ product (red bar). Finally, RISC departed or the guide's Alexa555 photobleached. The 3′ cleavage product containing the seed remains on the slide surface, allowing a new molecule of RISC to bind. (B) Rastergrams comparing 5′-tethered (426 individual molecules) and 3′-tethered (452 individual molecules) RNA targets fully complementary to let-7a.

To quantitatively assess the product release mechanism, we performed global fitting (Figure 3.6A) of a unified reaction scheme that accounts for all observed intermediates and products (Figure 3.6B). A global fit is possible for our kinetic data because the loss or production of different reaction species (e.g., 5′- and 3′-products) share one or more kinetic steps in the mechanism. By fitting multiple data sets simultaneously to the same kinetic mechanism, the rate constants for shared steps become global parameters constrained to be the same for all data sets. Our proposed reaction mechanism includes branched pathways for product release: one branch corresponds to the 5′ product being released first and the 3′ product released subsequently (Figure 3.5B, $k_{5'\,1st}$ followed by $k_{3'\,2nd}$), while in the other, the order of product release is reversed (Figure 3.5B, $k_{3'\,1st}$ followed by $k_{5'\,2nd}$). Both branches arrive at the same final state: two free products and free AGO2-RISC.

To account for the sigmoidal kinetics of product release (Figure 3.6A) we included an additional kinetic step. Because both branches of the mechanism share this step, its rate constant (*k*) was treated as global parameter for fitting. The rate constant (*k*) for the additional step likely corresponds to the rate of the slowest step in the target cleavage reaction—e.g., the conformational change in Argonaute that brings the catalytic Mg^{2+} near the scissile phosphate (Wang et al., 2009; Sheng et al., 2014). The global fit based on four experimentally measured product release curves obtained using 5′- and 3′-tethered targets

(Figures S3.8A and S3.8B) defined three of the five rate constants (k, $k_{5' 1st}$ and $k_{3' 1st}$). The rate constants ($k_{3' 2nd}$ and $k_{5' 2nd}$) for release of the 3′ or 5′ products following release of the other product were determined directly from the distributions of waiting times beginning with the departure of the first cleavage product and ending with the departure of the second (Figures S3.8A and S3.8B), after subtracting the photobleaching rate (Materials and Methods).

Figure 3.6

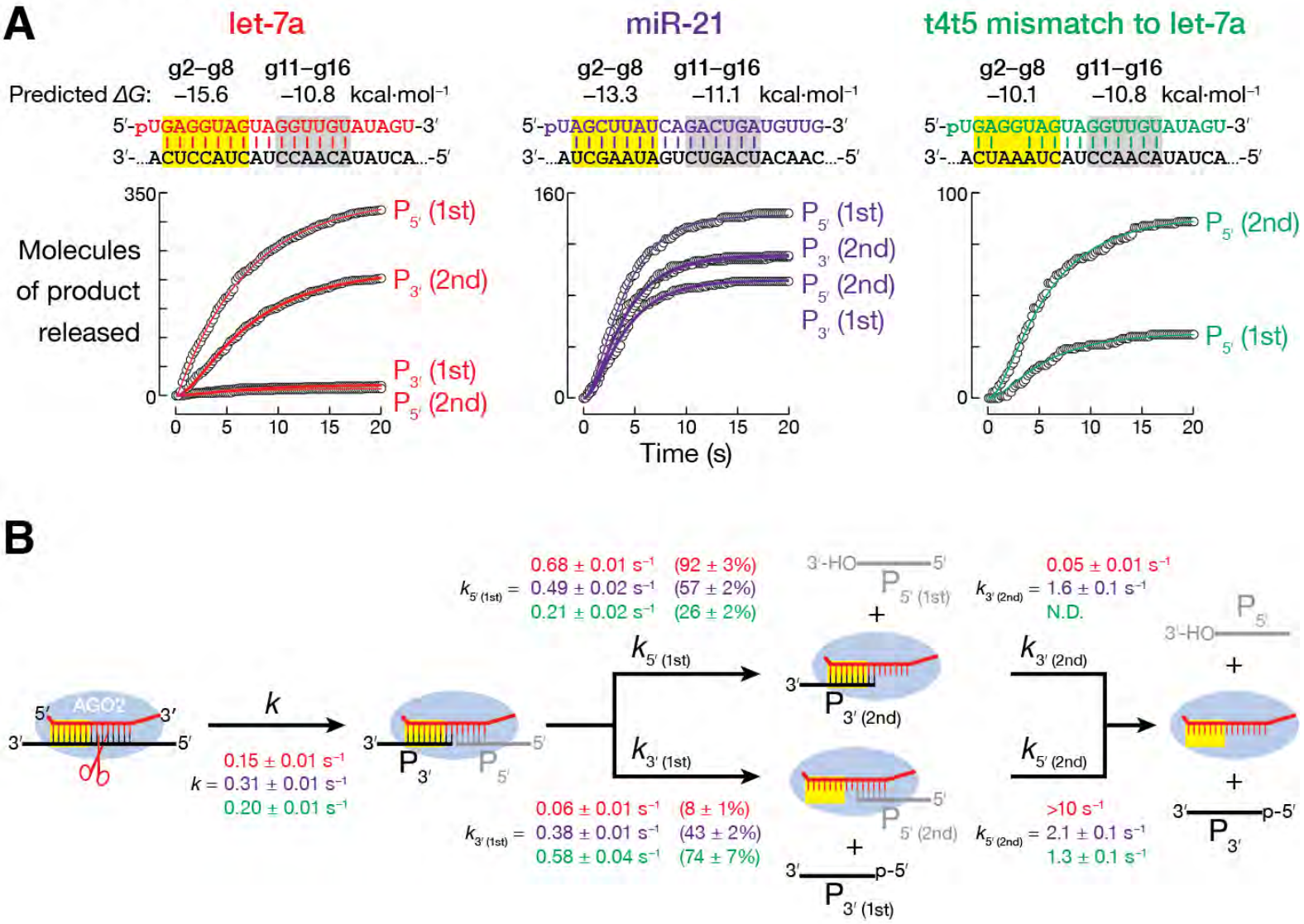

A
let-7a
miR-21
t4t5 mismatch to let-7a
g2–g8 g11–g16
Predicted ΔG: –15.6 –10.8 kcal·mol⁻¹
5′-pUGAGGUAGUAGGUUGUAUAGU-3′
3′-...ACUCCAUCAUCCAACAUAUCA...-5′
g2–g8 g11–g16
–13.3 –11.1 kcal·mol⁻¹
5′-pUAGCUUAUCAGACUGAUGUUG-3′
3′-...AUCGAAUAGUCUGACUACAAC...-5′
g2–g8 g11–g16
–10.1 –10.8 kcal·mol⁻¹
5′-pUGAGGUAGUAGGUUGUAUAGU-3′
3′-...ACUAAAUCAUCCAACAUAUCA...-5′
Molecules of product released
P5′ (1st)
P3′ (2nd)
P3′ (1st)
P5′ (2nd)
Time (s)
B
AGO2
k
0.15 ± 0.01 s⁻¹
k = 0.31 ± 0.01 s⁻¹
0.20 ± 0.01 s⁻¹
k5′ (1st) = 0.68 ± 0.01 s⁻¹ (92 ± 3%)
0.49 ± 0.02 s⁻¹ (57 ± 2%)
0.21 ± 0.02 s⁻¹ (26 ± 2%)
k3′ (1st) = 0.06 ± 0.01 s⁻¹ (8 ± 1%)
0.38 ± 0.01 s⁻¹ (43 ± 2%)
0.58 ± 0.04 s⁻¹ (74 ± 7%)
k3′ (2nd) = 0.05 ± 0.01 s⁻¹
1.6 ± 0.1 s⁻¹
N.D.
k5′ (2nd) = >10 s⁻¹
2.1 ± 0.1 s⁻¹
1.3 ± 0.1 s⁻¹

Figure 3.6: AGO2-Catalyzed Cleavage and Product Release

(A) Global fit analysis (Materials and Methods) of 5′- and 3′-tethered targets for AGO2 guided by let-7a or miR-21.
(B) The detailed kinetic scheme used for global fitting. Rate constants are color-coded according to (A).
Percentages in parentheses report the proportion of product molecules released first.

Seed Pairing Determines the Rate of Slicing and the Order of Product Release

To determine whether the features of cleavage and product release observed for let-7a RISC depend on the guide RNA identity and base-pairing stability with the target, we performed experiments paralleling those shown in Figure 3.5B but with 5′- and 3′-tethered targets fully complementary to miR-21 miRNA and AGO2-RISC loaded with miR-21. We also made a 5′-tethered let-7a target that contained mismatches with the let-7a seed at positions g4 and g5 (Figures S3.3 and S5C). We then carried out global fitting of the kinetic scheme in Figures 3.6B and S3.8A to these data to determine slicing and product release rates, as well as the order of product release.

Slicing rate depended on guide strand identity: let-7a has the slowest slicing rate we measured (k = 0.15 s^{-1}), while miR-21 AGO2-RISC cleaves its target twice as fast (k = 0.31 s^{-1}). The let-7a has the strongest seed pairing with $\Delta G = -15.6$ kcal·mol^{-1}. The miR-21 seed is significantly weaker: $\Delta G = -13.3$ kcal·mol^{-1}. The let-7a target with seed mismatches that weaken the seed base pairing by about 5 kcal·mol^{-1} ($\Delta G = -10.1$ kcal·mol^{-1}) is cleaved faster than the fully complementary let-7a target (k = 0.20 s^{-1}), but still slower than miR-21. This trend appears to inversely correlate with the stability of seed pairing to the target, albeit not precisely, indicating that there may be additional determinants of the cleavage rate (Figure S3.8C). Interestingly, miR-21 AGO2-RISC failed to

cleave or even detectably bind the miR-21 target that contains g4g5 mismatches (data not shown), indicating that weakening seed base pairing below a certain threshold prevents target recognition and cleavage.

The stability of seed pairing with respect to the stability of base pairs in the 3′-part of the guide strand determines the order of product release, as clearly evidenced by the proportion of the reaction directed through one of the two product release branches (Figure 3.5B, values in parentheses). For let-7a, whose seed pairing (g2–g8; $\Delta G = -15.6$ kcal·mol^{-1}) is predicted to be more stable than the pairing of the 3′-half of the guide with the 5′ product (g11–g16; $\Delta G = -10.8$ kcal·mol^{-1}), the 5′ product was released before the seed-matching 3′ cleavage product for 92% of molecules.

For let-7a, release of the 3′ cleavage product from RISC limits the rate of enzyme turnover, k_{cat}. Is 3′ product release generally rate-determining for RISC turnover? To test this, we analyzed two additional miRNA:target combinations: miR-21 paired to a fully complementary target and let-7a paired to a target bearing two mismatches to the let-7 seed. Unlike let-7a, miR-21 has a smaller predicted difference in the stabilities of target pairing to the seed and 3′-half ($\Delta G_{g2-g8} = -13.3$ kcal·mol^{-1} versus $\Delta G_{g11-g16} = -11.1$ kcal·mol^{-1}). For miR-21, only 57% of the 5′ product was released first. For let-7a paired to a target mismatched to guide positions g4g5, the 5′ product is predicted to pair with let-7a more stably than the partially seed-matching 3′ product

($\Delta G_{g2\text{-}g8} = -10.1$ kcal·mol^{-1} versus $\Delta G_{g11\text{–}g16} = -10.8$ kcal·mol^{-1}). For this guide:target pair, only 26% of the 5′ product was released before the 3′ cleavage product. We conclude that the course of product release—5′ first or 3′ first—does not follow a strict order but rather reflects the sequences of the guide and target.

Product release rates followed the base pair stability trends. For the 3′ product, the release rate of let-7a paired to a fully complementary target (0.06 s^{-1}) was tenfold slower than for let-7a paired to the g4g5:t4t5 mismatched target (0.6 s^{-1}), reflecting their different predicted free energies of base pairing between the seed sequence and the 3′ product (–15.6 kcal·mol^{-1} versus –10.1 kcal·mol^{-1}). For the 5′ product, the release rates were more similar: 0.7 s^{-1} for let-7a, 0.5 s^{-1} for miR-21, and 0.2 s^{-1} for let-7a with the seed-mismatched target. The free energies of pairing for all three 5′ products were roughly –11 kcal·mol^{-1}.

Release of the First Product Promotes Release of the Second

Formally, there are two rates for the release of each cleavage product: a rate for when the product departs first and a rate for when the same product departs second, the other product having already dissociated. For example, $k_{5'\,1st}$ is the rate for 5′ product release in the presence of bound 3′ product, while $k_{5'\,2nd}$ is the rate for 5′ product release after the 3′ product has already left. Our data suggest that release of the first product promotes release of the second product (Figure 3.6). For example, the rates of the 5′ and 3′ products of miR-21 were both ~4-

fold faster when they were released second rather than first. Similarly, release of the 5′ product of cleavage of the let-7a seed-mismatched target was $0.21 \pm 0.01\ s^{-1}$ when released first, but $1.3 \pm 0.1\ s^{-1}$ when released second. A notable exception was the seed-matched 3′ product of let-7a target, by far the most stably bound product we examined. This 3′ cleavage product dissociated at ~$0.05\ s^{-1}$ regardless of the presence of the 5′ product.

We can imagine two mechanisms by which the departure of one product can accelerate dissociation of the other. Stacking interactions between the terminal bases of the products may mutually stabilize their binding. Supporting this view, the 5′ product of miR-21 leaves ~4-fold faster when it departs after the 3′ product (and vice versa)—an ~$0.9\ kcal \cdot mol^{-1}$ difference in stability. This difference in ΔG is within the range of the contributions of dangling nucleotides to RNA helix stability (Xia et al., 2001). Alternatively, departure of one of the two products may facilitate a conformational change that destabilizes the second product. Such a conformational change might correspond to the return of the endonuclease active site to the conformation present prior to zippering of the guide:target helix 3′ to the seed sequence (Wang et al., 2009; Elkayam et al., 2012; Schirle and MacRae, 2012; Faehnle et al., 2013).

Strong Seed Pairing Slows RISC Turnover

The rates for target cleavage and product release allow calculation of the overall turnover rate: $k_{cat} = k \cdot k_{5' 1st} \cdot k_{3' 2nd}/(k \cdot k_{5' 1st} + k \cdot k_{3' 2nd} + k_{5' 1st} \cdot k_{3' 2nd})$ when the

5′ product is released first, and $k_{cat} = k \cdot k_{3'\,1st} \cdot k_{5'\,2nd}/(k \cdot k_{3'\,1st} + k \cdot k_{5'\,2nd} + k_{3'\,1st} \cdot k_{5'\,2nd})$ when the 3′ product is released first. The two pathways result in similar k_{cat} values. For let-7a, the calculated k_{cat} value, 0.036 ± 0.002 s^{-1}, agrees well with k_{cat} determined by traditional, ensemble initial velocity analysis (0.066 ± 0.004 s^{-1}; Figure 3.1C). The calculated turnover rate was about fourfold faster for both miR-21 (0.16 ± 0.1 s^{-1}) and let-7a with the g4g5 seed-mismatched target (0.13 ± 0.1 s^{-1}). The slower k_{cat} for let-7a reflects the stronger seed pairing to its fully complementary target ($\Delta G = -15.6$ kcal·mol^{-1}, $k_{3'\,1st} = 0.06$ s^{-1}, $k_{3'\,2nd} = 0.05$ s^{-1}). This slow 3′ product release step for let-7a limits the overall turnover rate. Both miR-21, with its weaker pairing strength in its seed, and let-7a paired to the target whose seed pairing was intentionally weakened with g4g5 mismatches, direct faster cleavage than let-7a with a fully complementary target, because their product release rates are comparable to or faster than k, the apparent RISC cleavage rate (Figure 3.6B). Thus, guide RNAs, including siRNAs, with more stable base pairing between their seed and their target are predicted to cleave fewer targets per unit time than targets whose rate of 3′ product release is not rate-determining.

The autoantigen protein La has been proposed to facilitate multiple-turnover target cleavage by human AGO2 RISC (Liu et al., 2011). Thus, La protein might accelerate release of the cleaved target from RISC, so that product release is no longer rate-determining in vivo (Figure 3.7A). To test this

idea, we measured the rate of multiple-turnover target cleavage by let-7a-AGO2 RISC in the absence or presence of an equimolar (1 nM) or excess (50 or 500 nM) amount of La protein (Figure 3.7B). We were unable to detect any difference in the rate of target cleavage among the four reaction conditions. We conclude that, in cells, La protein is unlikely to overcome the rate-limiting step of product release for let-7a-AGO2 RISC.

Figure 3.7

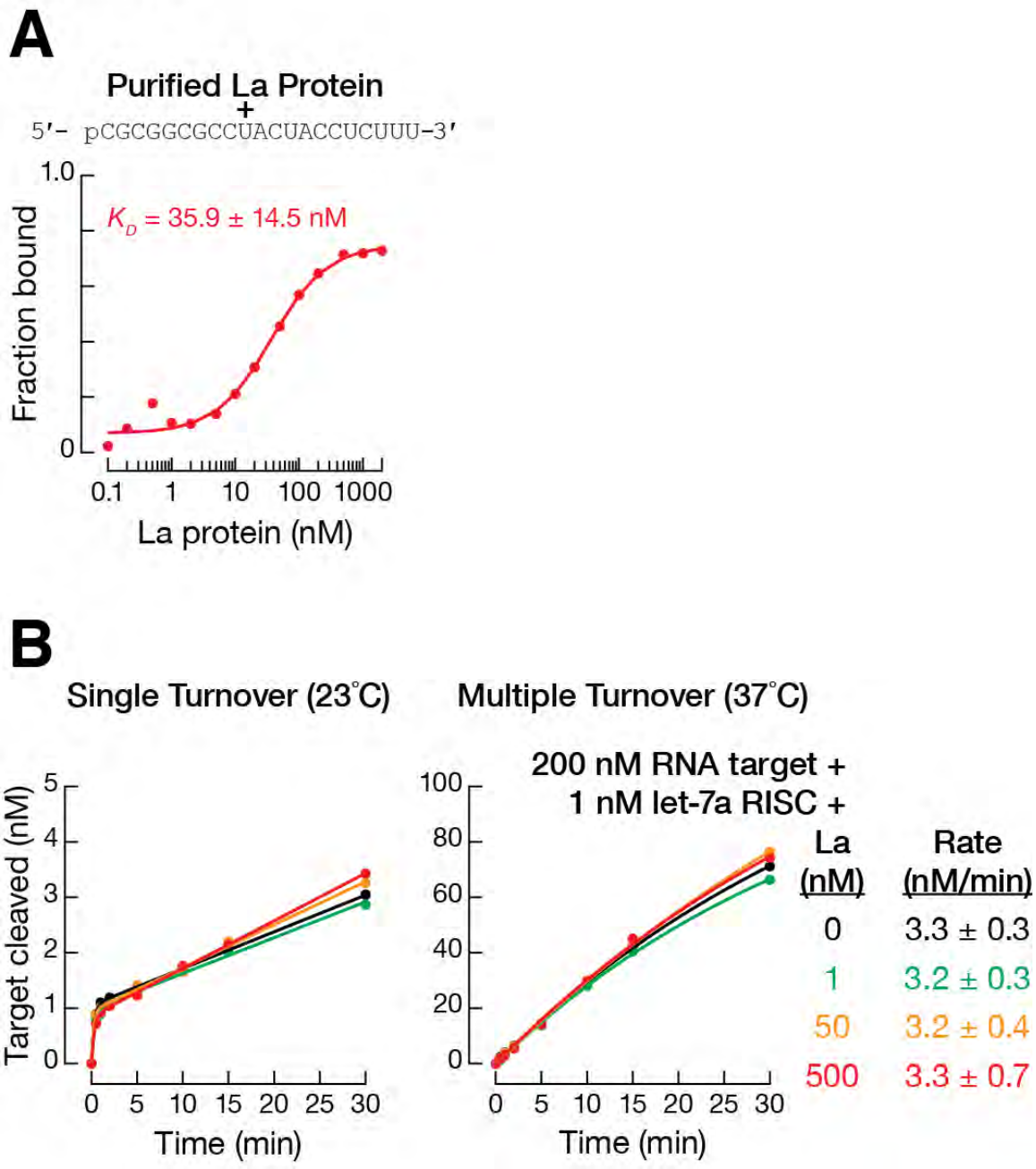

A
Purified La Protein
+
5′- pCGCGGCGCCUACUACCUCUUU-3′
1.0
0
Fraction bound
K_D = 35.9 ± 14.5 nM
0.1
1
10
100
1000
La protein (nM)
B
Single Turnover (23°C)
Multiple Turnover (37°C)
200 nM RNA target +
1 nM let-7a RISC +
Target cleaved (nM)
Time (min)
La (nM)
Rate (nM/min)
0
3.3 ± 0.3
1
3.2 ± 0.3
50
3.2 ± 0.4
500
3.3 ± 0.7

Figure 3.7: Measurement of the RNA binding protein La on multiple turnover activity

(A) Dissociation constant determination for purified La determined by gel-shift analysis (Materials and Methods). (B) RISC target cleavage under conditions of target excess. Target cleavage was performed at 23˚C and 37˚C (Materials and Methods).

AGO2 Distinguishes between miRNA-Like Binding Sites and Cleavage Products

Although quantitatively different from the energetics of nucleic acid hybridization, the free energy of base pairing between the seed sequence and its target influences the rates of all steps in the RNAi reaction, including binding and dissociation of RISC, cleavage of the target, and release of the cleaved products. However, one aspect of RISC function emerges from our studies that is not predicted by the stability of guide-target base pairing: AGO2 appears to discriminate between a miRNA-like binding site, which typically pairs only with nucleotides g2–g8, and binding to the seed-matched, 3′ product of target cleavage, which pairs with nucleotides g2–g10 (Figure 3.8).

We measured the k_{off} for a full-length target RNA bound to let-7a-guided AGO2-RISC for a target complementary only to the let-7a seed (g2–g8) and for targets with one (g9:t9) or two (g9g10:t9t10) additional base pairs beyond the seed, as would be found in the 3′ product after target cleavage. The standard rules for nucleic acid hybridization predict that the addition of one or two additional base pairs beyond the seed should decrease ΔG, slowing dissociation of RISC (Figure 3.8). Counterintuitively, these additional base pairs increased the dissociation rate ~4-fold: k_{off} was 0.014 ± 0.001 s^{-1} when g2–g9 paired with the target and 0.015 ± 0.001 s^{-1} (τ ~66 s) when g2–g10 paired, but 0.0036 s^{-1} (τ ~280 s) when only g2–g8 were paired. In vivo, AGO2 has little opportunity to

bind targets with complementarity from t2–t10, but this pairing scheme is typically found in the 3′ cleavage product generated by the RNAi pathway. RISC departed >3-fold faster from the 3′ product generated by target cleavage (k_{off} = 0.05 ± 0.01 s^{-1}) than from the full-length target complementary to g2–g10, and ~14-fold faster than from the full-length, seed-matched (t2–t8) target. When the 3′ cleavage product was subsequently bound by other RISC molecules, they depart at essentially the same rate as the RISC that first catalyzed cleavage, k_{off} = 0.044 ± 0.001 s^{-1} (τ ~23 s; Figure 3.8B).

Figure 3.8

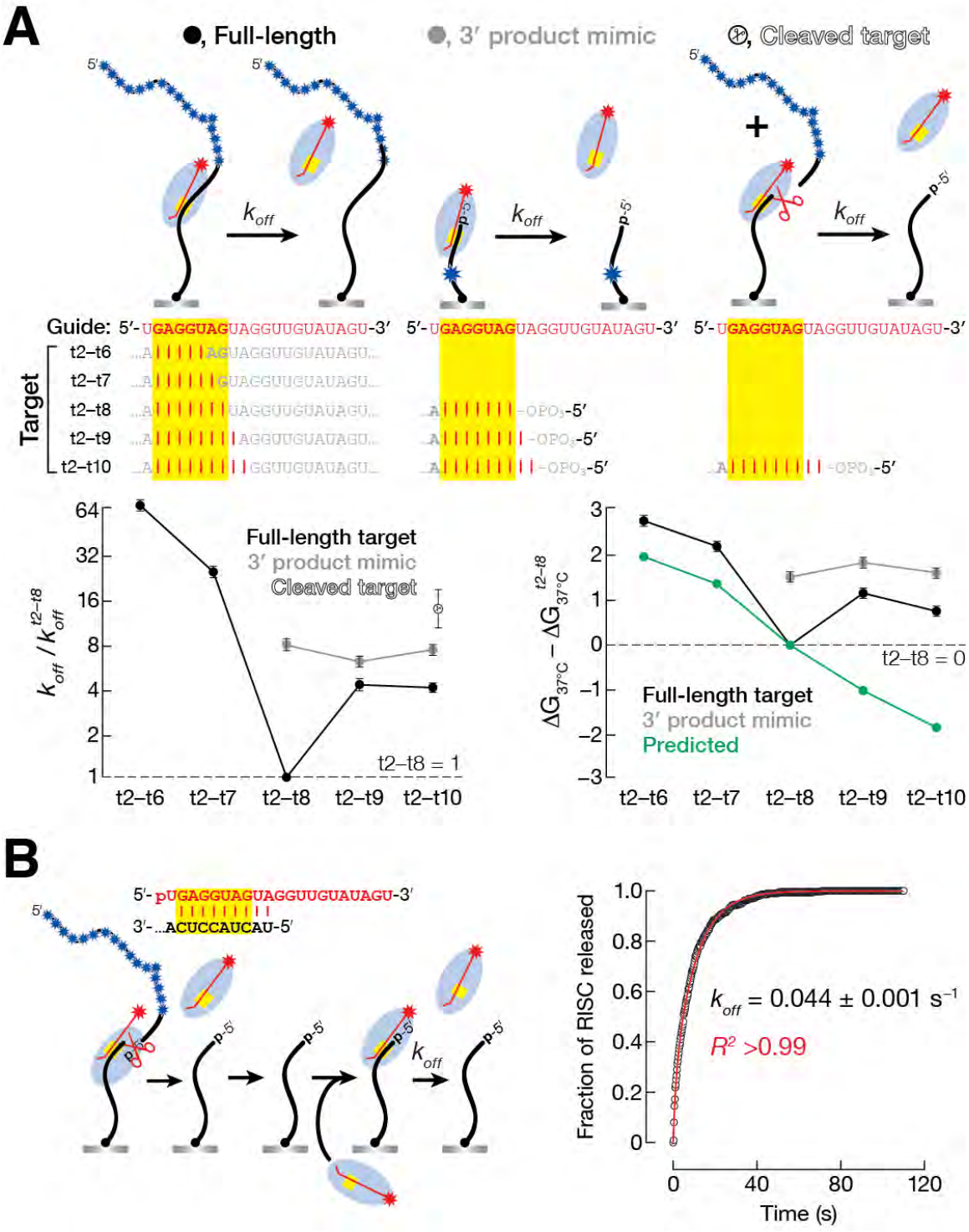

A
●, Full-length
●, 3′ product mimic
⊛, Cleaved target
Guide: 5′-UGAGGUAGUAGGUUGUAUAGU-3′
Target
t2–t6
t2–t7
t2–t8
t2–t9
t2–t10
Full-length target
3′ product mimic
Cleaved target
t2–t8 = 1
Predicted
t2–t8 = 0
B
5′- pUGAGGUAGUAGGUUGUAUAGU-3′
3′-...ACUCCAUCAU-5′
Fraction of RISC released
k_{off} = 0.044 ± 0.001 s^{-1}
R^2 >0.99
Time (s)

Figure 3.8: Argonaute can Distinguish between miRNA Targets and Cleaved Products

(A) Effects of additional complementarity and 3′ target length on RISC binding and dissociation. k_{off} was measured directly, correcting for photobleaching, except for the seed-matched, full-length target, whose dissociation was slower than the rate of photobleaching and was therefore measured by varying laser exposure time and extrapolating to no illumination. $\Delta\Delta G_{37°C}$ was calculated from K_D (= k_{off} / k_{on}); theoretical $\Delta\Delta G_{37°C}$ was predicted using nearest-neighbor analysis to estimate ΔG. (B) Experimentally measured k_{off}, corrected for photobleaching, for let-7a-guided mouse AGO2-RISC dissociating from the 3′ product of a previously cleaved, 3′-tethered, fully complementary target RNA.

One potential explanation for the accelerated departure of RISC from a 3′ cleavage product compared to a seed-matched target is that the seed-matched target extends far beyond position t10, providing greater opportunities for non-sequence specific interactions between AGO2 and its target. To test this idea, we designed a series of 5′ monophosphorylated, 30 nt RNAs that end at either position t8, t9, or t10 so as to mimic a 3′ cleavage product (Figure 3.8A). Compared to the full-length target complementary to guide positions g2–g10 (k_{off} = 0.014 ± 0.001 s^{-1}), the 3′ cleavage product mimic dissociated faster (k_{off} = 0.026 ± 0.001 s^{-1}; Figure 3.8), further supporting the idea that AGO2 makes sequence-independent contacts with its RNA targets (Ameres et al., 2007). RISC also departed approximately eight times faster from a seed-matched (t2–t8, k_{off} = 0.028 ± 0.001 s^{-1}) mimic of the 3′ cleavage product than from the seed-matched, full-length target.

Nonetheless, such contacts do not explain why RISC departed faster when paired with t2–t10 than with only t2–t8 of a full-length target. If non-sequence-specific protein-RNA interactions do not explain the faster departure from the 3′ cleavage product than for a seed matched target, how does g9g10:t9t10 pairing alter the properties of RISC so that it binds more weakly to an RNA with nine potential base pairs than to an RNA with which it makes only seven? Extending base pairing beyond position g8 of the seed likely requires an energetically unfavorable conformational rearrangement in RISC.

The structure of human AGO2 with a guide RNA complementary to nucleotides t2–t9 of short target RNA suggests that pairing beyond t8 requires opening the central cleft of AGO2 (Schirle et al., 2014). Such a presumably disfavorable change in conformation would explain why base pairing beyond guide position g8 is atypical for miRNA-target interactions. These properties of AGO2 were not anticipated from thermodynamic predictions for the strength of nucleic acid base pairing (Figures 3.8 and S3.8).

DISCUSSION

Our data demonstrate directly that Argonaute proteins from both bacteria and mammals accelerate the rate at which their guide strands find complementary targets. Argonautes can accelerate on-rates as much as 250 times, and both TtAgo, which is DNA-guided, and mouse AGO2, which is RNA-guided, can enhance target finding for both DNA and RNA targets. The acceleration of target binding by AGO2 requires seed complementarity with its target, consistent with pre-organization of the seed sequence by Argonautes playing a major role in this phenomenon. However, seed nucleotides do not contribute equally to target binding: g2–g5:t2–t5 base pairs contribute more to the initial binding of RISC to target RNA than g6–g8:t6–t8 base pairs, which function mainly to slow dissociation of RISC from its target after a successful encounter. Remarkably, the existence of seed sequence subdomains with distinct functions in target

binding was first predicted from the structure of human AGO2-RISC bound to a seed-matched target RNA (Schirle et al., 2014).

Because target binding is a bimolecular process, the rate at which RISC finds its targets in vivo depends on the concentrations of both target sites and Argonaute-bound guide. In light of recent reports that in vivo target site concentration is quite high and exceeds that of miRNA-guided RISC (Denzler et al., 2014), the rate of target finding is unlikely to limit the speed at which miRNAs destabilize their target mRNAs. Furthermore, the rates of target binding in our single-molecule experiments approach the speed of macromolecular diffusion, suggesting that additional proteins such as poly(A) binding protein (PABP) cannot promote association of RISC with seed-complementary targets, as has been previously suggested (Moretti et al., 2012). Of course, our studies were conducted with purified components outside of the cell, and future experiments measuring the diffusion and rate of target binding of bacterial and mammalian RISCs in living cells will be required to test these ideas.

In addition to accelerating hybridization of guides to targets, Argonaute proteins alter how quickly target sequences dissociate from RISC. For mouse AGO2, k_{off} can be as slow as 0.003 s^{-1}. That is, at 37°C, more than 5 minutes is required for an average RISC to depart from a typical miRNA-binding site. In fact, many miRNA-guided Argonautes are predicted to bind their target sites with greater affinity than most known RNA-binding proteins. Consequently,

RISC likely serves to stabilize the binding of other proteins, such as GW182 or mRNA-degrading nucleases, to the targets of miRNAs. The high binding affinity of RISC derives mainly from its slow dissociation rate, which is far slower than would be expected for seven base pairs between RNA strands: the K_D of a guide RNA in RISC bound to a seed-match target is ~4 million times tighter and dissociates ~14,000 times slower than for RNA alone.

Many of these properties spring from the unique configuration of guide nucleotides g2–g8 within Argonaute proteins, which pre-organize them into the seed sequence. Our data show that the sequence of the seed, whose existence was detected by insight or computation before the first structure of an Argonaute-bound guide was available (Lai and Posakony, 1998; Lai, 2002; Lewis et al., 2003), influences the rate of dissociation, with RISC dissociating from its seed-matched target far more slowly when guided by let-7a than by miR-21. Moreover, mismatches in the seed increase the dissociation of RISC from its otherwise fully complementary seed-matched target more than what would be predicted from the thermodynamics of base pairing (Figure S3.7). For example, a g2g3:t2t3 mismatch increased k_{off} ~3,000-fold, an increase in free energy ~6 kcal·mol^{-1}; for RNAs hybridizing in the absence of protein, the predicted change was just ~1 kcal·mol^{-1}. Together, our data suggest that Argonautes do more than simply pre-organize the seed—which would pre-pay the entropic cost of hybridization to a similar extent for all seed sequences (Figure 3.9). Argonaute

proteins must also provide an environment in which the free energy of each non-seed base pair is decreased, but the rank order of base pairing strength predicted by nearest-neighbor analysis remains essentially unchanged. Guide positions g9 and g10 are a notable exception to this idea. t9 and t10 target complementarity to the guide provide no increase in binding affinity, and, when not at the 5′ end of the target RNA (a cleavage product), actively destabilize RISC binding. Thus, miRNA prediction algorithms that penalize or at least give no credit for pairing at nucleotides t9 and t10 are more likely to identify biologically important mRNA targets.

Intriguingly, mismatches at the 5′ end of the seed have the greatest effect on k_{off} (Figure 3.3), even when these mismatches are not predicted to have the greatest effect on base pairing stability (Figure S3.7). These data suggest that even within the seed sequence, Argonaute assigns greater value to mismatches before guide position g5 than from g5–g8. In fact, the structure of AGO2, loaded with a guide and bound to a seed matched target, suggests that pairing beyond g5:t5 requires a conformational rearrangement in the protein (Schirle et al., 2014). Such a rearrangement could account for the reduced contribution of bases g5–g8 to binding affinity.

Although most miRNA-binding sites match the seed sequence, some sites use centered pairing, ≥11 nt of contiguous complementarity to the guide typically starting at position t4 (Shin et al., 2010). While target mismatches with

the 5′ end of the seed normally prevent stable miRNA binding, extensive complementarity to the center and 3′ end of a miRNA may provide compensatory binding energy. Supporting this idea, let-7a-RISC binding to a seed-matched target containing a g4g5:t4t5 mismatch was rescued by additional base pairs 3′ to the seed (Figures 3.3 and S3.5). Regulation of the *Glutathione S-transferase Mu 3* (*GSTM3*) mRNA by miR-21, which pairs with nucleotides t5–t16 provides an example of a centered site (Shin et al., 2010). We speculate that the faster than typical off-rate of the miR-21 seed (Figures 3.3 and S3.5) facilitated the evolution of this type of site. Because centered sites extend well into the 3′ region of the guide, centered pairing may offset the destabilizing effects of g9g10:t9t10 complementarity pairing conformation. Alternatively, the absence of g2–g5:t2–t5 seed pairing may alleviate the conformational strain associated with g2–g10:t2–t10 complementarity.

The extent of complementarity between guide and target determines whether RISC simply binds or binds and cleaves. In animals, miRNAs typically guide binding, whereas siRNAs usually direct target cleavage. Our data suggest that mammalian AGO2 can discriminate between seed-matched “miRNA targets,” to which RISC remains tightly bound, and “siRNA targets,” with which it pairs more extensively. After target cleavage, release of the two cleavage products follows no strict order, and AGO2 creates an environment in which dissociation of either of the cleavage products accelerates departure of the

other product. Thus, release of either cleavage product can be rate-determining, but is nevertheless faster than would be predicted from the extent of base pairing with the guide. For the 3′ product, its faster than expected release can be explained, in part, by the loss of non-sequence specific protein-RNA contacts after the 5′ product leaves.

TtAgo also senses the type of target to which it is bound. In vivo, TtAgo catalyzes cleavage of foreign DNA with extensive complementarity to its 16 nt DNA guide (Swarts et al., 2014a). Consistent with this idea, TtAgo dissociated slowly from a fully complementary target DNA, but quickly from a target that was only partially complementary ($\tau \leq 1$ s at 37°C; Table 3.1). Unlike mouse AGO2, which appears to have evolved to bind stably only to RNA, TtAgo readily binds and cleaves both DNA and RNA. In contrast, mouse AGO2 finds complementary sites in DNA more rapidly than RNA but fails to bind stably to DNA ($\tau \leq 2$ s ; Table 3.1).

Our efforts to apply single-molecule methods to study the fundamental properties of bacterial and mammalian Argonaute proteins demonstrate the utility of this approach. Unlike classical ensemble approaches for measuring the kinetics of RNA binding and cleavage by RISC, single-molecule studies allow direct and continuous observation of rapid events for both the target and guide. We anticipate that future studies will extend this approach to other Argonaute proteins, including the animal-specific PIWI clade, which defends the germline

against transposons, and plant Argonautes, which mediate both mRNA cleavage (Llave et al., 2002; Tang et al., 2003) and repression of mRNA translation (Brodersen et al., 2008; Iwakawa and Tomari, 2013), as well as to more complex sets of proteins that collaborate with Argonautes to carryout the repression of gene expression.

Figure 3.9

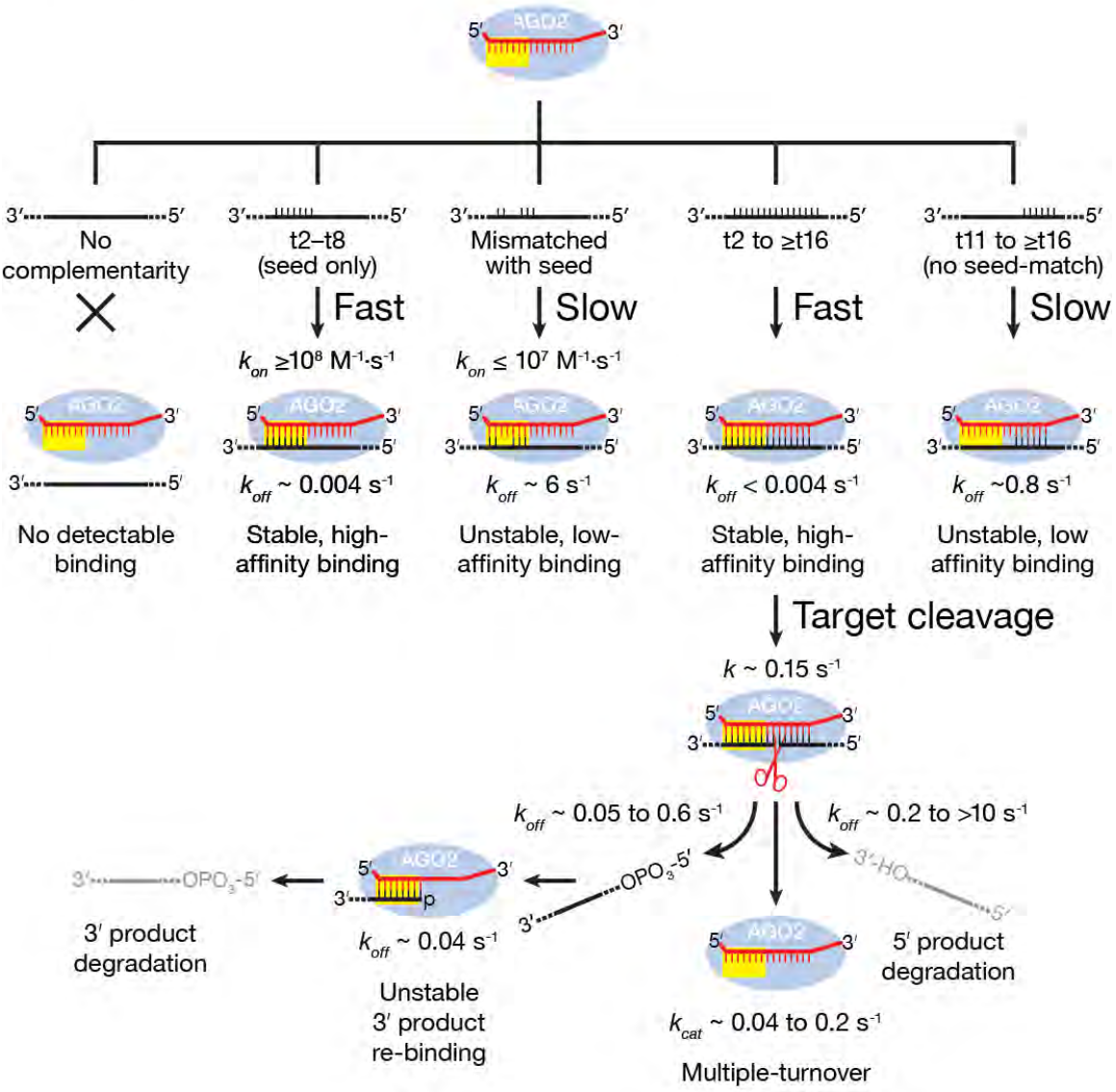

AGO2
No complementarity
t2–t8 (seed only)
Mismatched with seed
t2 to ≥t16
t11 to ≥t16 (no seed-match)
Fast
Slow
Fast
Slow
k_{on} ≥10⁸ M⁻¹·s⁻¹
k_{on} ≤ 10⁷ M⁻¹·s⁻¹
k_{off} ~ 0.004 s⁻¹
k_{off} ~ 6 s⁻¹
k_{off} < 0.004 s⁻¹
k_{off} ~0.8 s⁻¹
No detectable binding
Stable, high-affinity binding
Unstable, low-affinity binding
Stable, high-affinity binding
Unstable, low affinity binding
Target cleavage
k ~ 0.15 s⁻¹
k_{off} ~ 0.05 to 0.6 s⁻¹
k_{off} ~ 0.2 to >10 s⁻¹
3′ product degradation
k_{off} ~ 0.04 s⁻¹
Unstable 3′ product re-binding
5′ product degradation
k_{cat} ~ 0.04 to 0.2 s⁻¹
Multiple-turnover

Figure 3.9: A Kinetic Model for Mouse AGO2-RISC Function

Materials and Methods

Preparation of siRNAs and RNA Targets

Synthetic RNAs for passenger and guide strands (GE Healthcare Dharmacon, Lafayette, CO; Sigma-Aldrich Corp, St. Louis, MO) were deprotected and gel purified (Table S3.2). Guide RNAs and targets bearing a 3′ amine modification or 3′ product mimic targets containing an internal 2′-amino uridine were reacted with Alexa Fluor 555 or Alexa Fluor 647 *N*-hydroxysuccinimide ester (Life Technologies, Thermo Fisher Scientific, Grand Island, NY) in 0.1M sodium tetraborate (pH 8.0) overnight at 4°C and then purified from a denaturing 20% polyacrylamide gel and eluted overnight from the gel slice into 0.4 M NaCl, 25 mM EDTA rotating in the dark at room temperature, and then precipitated with ethanol. (The Alexa dye caused an ~1 nt increase in apparent mobility.) All target RNAs were transcribed using T7 RNA polymerase and DNA templates generated by PCR (Table S3.2) from pGL2-Control vector (Promega, Madison, WI) and then gel purified (Haley et al., 2003). RNA substrates for kinetic analyses were in vitro transcribed with ^{32}P α-UTP (6,000 Ci/mmol; Perkin Elmer, Waltham, MA) and then capped with ^{32}P α-GTP (3,000 Ci/mmol; Perkin Elmer) and guanylyl transferase as described (Wee et al., 2012). All RNAs were gel purified and concentrations were determined by absorbance at 260 nm prior to addition of Alexa dye.

RNA targets tethered via the 5′ end were transcribed in the presence of 5′-biotin-GMP (TriLink Biotechnologies, San Diego, CA) at a ratio of 4:1 relative to GTP, then gel purified. The 3′ ends of these RNAs were trimmed using a "gapmer" antisense oligonucleotide (Table S3.2) and RNaseH (Life Technologies). The trimmed RNA was gel purified and concentration was determined by absorbance at 260 nm. Alexa Fluor 647 dyes were added to the 3′ end of the in vitro transcribed, 5′-biotinylated, RNase H-trimmed RNA by template-directed extension using Klenow polymerase fragment. In a typical labeling procedure, 50–500 pmol RNA target was annealed to a 1.5–2-fold molar excess of DNA template oligonucleotide (Table S3.2) in 10–20 μL 10 mM HEPES-KOH (pH 7.4), 20 mM NaCl and 0.1 mM EDTA at 70°C followed by slowly cooling to 30°C over 20 min. Afterwards, the annealed strands (< 25% of final reaction volume) were added without further purification to the 3′ extension reaction, comprising 1× NEB buffer 2 (New England Biolabs, Ipswich, MA), 0.5 mM dATP, 0.5 mM dCTP, 0.1 mM or a 2-fold molar excess of Alexa Fluor 647-aminohexylacrylamido-dUTP (Life Technologies) over the theoretical amount of dUTP required for Klenow extension, whichever was greater, and 0.25 U/μL Klenow fragment (3′→5′) exo-minus, New England Biolabs) and incubated at 37°C for 1 h. The reaction was quenched by the addition of 500 mM (f.c.) ammonium acetate and 20 mM (f.c.) EDTA. A 10-fold molar excess of "Trap" oligonucleotide (Table S3.2) was added to the DNA

template oligonucleotide. The entire reaction was precipitated overnight at −20°C in 3 volumes of ethanol. The labeled target was recovered by centrifugation, dried, and dissolved in Gel Loading Buffer II (Life Technologies AM8547), incubated at 95°C for 60 seconds, and gel purified.

RNA targets tethered by their 3′ ends were in vitro transcribed as described above except 5′-biotin-GMP was omitted. After gel purification, the 5′ triphosphate was converted to a monophosphate using RNA 5′ polyphosphatase (Epicentre, Madison, WI), and then the 3′ end was trimmed using a gapmer antisense oligonucleotide (Table S3.2) and RNaseH (Life Technologies). To label the 5′ end of the target, a DNA extension was transcribed in the presence of Alexa Fluor 647-aminohexylacrylamido-dUTP using a DNA oligonucleotide template and Klenow polymerase. The labeled Alexa647-dUTP DNA was gel purified, and its concentration was determined by absorbance at 260 nm. To ligate DNA to the 5′ end of a RNA, we first ligated a synthetic DNA/RNA linker to the 3′ end of the gel-purified DNA made by Klenow extension using a synthetic DNA splint and T4 DNA ligase (Table S3.2). The Alexa 647-containing DNA bearing 30 nt of RNA at its 3′ end was ligated onto the 5′ end of the in vitro transcribed, 5′ monophosphorylated target as described (Moore and Query, 2000). The 5′ DNA with Alexa647-RNA target was gel purified and a 3′ biotin was added by Klenow extension using the template oligonucleotide for the synthesis of a short 3′ DNA extension with a single biotin

(Table S3.2) as described above, except that the Klenow reaction contained 0.5 mM dUTP instead of Alexa647-dUTP, and 0.1 mM biotin-dCTP (TriLink Biotechnologies) instead of dCTP.

RISC Purification and Ensemble Kinetics

S100 extract was generated from SV40 large T-antigen-immortalized *Ago2–/–* MEFs that stably over-express mouse AGO2 (O'Carroll et al., 2007). Cell extract was prepared as described (Dignam et al., 1983) except that acetate, not chloride, was used as the counter ion (Wee et al., 2012; Flores-Jasso et al., 2013). To load AGO2-RISC, 25 nM let-7a or miR-21 siRNA duplex was incubated in S100 extract (50% total reaction volume) for 1.5 h at 37˚C in 15 mM HEPES-KOH (pH 7.9) 100 mM potassium acetate, 5 mM magnesium acetate, 5 mM dithiothreitol, 1 mM ATP, 25 mM creatine phosphate, 30 $\mu g \cdot mL^{-1}$ creatine kinase. RISC was purified as described (Flores-Jasso et al., 2013) except that excess competitor oligo (Table S3.2) was removed by incubating the eluate with streptavidin paramagnetic beads (Dynabeads MyOne Streptavidin T1 10 $mg \cdot mL^{-1}$, Life Technologies) for 15 min at 4˚C.

Briefly, the assembled AGO2-RISC was incubated overnight at 4˚C a biotinylated, 2′-*O*-methyl capture oligo linked to paramagnetic beads (Dynabeads MyOne Streptavidin T1). The eluted RISC was concentrated, and the potassium acetate concentration adjusted to 100 mM (f.c.) by centrifugal ultrafiltration (Amicon Ultra-centrifugal filter, 10K MWCO, EMD Millipore,

Billerica, MA). The concentration of active, purified RISC was measured by pre-steady-state target cleavage assays at 23˚C in the presence of 100 or 200 nM ^{32}P-radiolabeled target RNA. For Michaelis-Menten kinetics, cleavage assays were performed at 25˚C with 0.1 nM purified let-7a-RISC and 0.1 to 5 nM ^{32}P-radiolabeled target RNA. For experiments with La protein, cleavage assays were performed at 23˚C or 37˚C with 1 nM purified let-7a-RISC, 200 nM ^{32}P-radiolabeled target RNA, and 0–500 nM La purified from bovine thymus (Arotec Diagnostics, Petone, Wellington, New Zealand).

La RNA-Binding Assays

To establish that La was active, the K_D of La binding RNA was determined by native-gel mobility analysis as described (Teplova et al., 2006) except that 1.5 nM ^{32}P-radiolabeled, 21 nt RNA oligonucleotide (Table S3.2, Sigma-Aldrich Corp) containing three 3′ terminal uridines was incubated with 0–2 μM purified La (Arotec Diagnostics) for 10 min at 37˚C in 50 mM Tris-HCl, pH 7.9, 50 mM potassium acetate, 5 mM magnesium acetate, 1 mM dithiothreitol, 10% (w/v) glycerol, 0.01% (v/v) Igepal CA-630. Samples were resolved by electrophoresis through an 8% native, polyacrylamide gel run at 4˚C for 10 min at 250 V and then at 100 V for 1 h. The gel was dried, exposed to an image plate, and then scanned and analyzed using a FLA-7000 laser scanner (GE Healthcare Bioscience) and MultiGuage 3.0 software (Fujifilm, Tokyo). The K_D was

determined using Igor Pro 6.36 (WaveMetrics, Lake Oswego, OR) by fitting the binding data to

$$f = \frac{[E_T] + [S_T] + [K_D] - \sqrt{-([E_T] + [S_T] + [K_D]) \cdot -([E_T] + [S_T] + [K_D]) - (4 \cdot ([E_T] \cdot [S_T])}}{2 \cdot [S_T]}$$

where *f* is fraction target bound, $[E_T]$ is total enzyme concentration, $[S_T]$ is total RNA target concentration, and K_D is the apparent equilibrium dissociation constant.

***Thermus thermophilus* Argonaute Expression, Assembly, and Purification**

Expression and purification of TtAgo was as described (Wang et al., 2008b) except for the final chromatography step. Briefly, TtAgo was amplified from genomic DNA and cloned into pET SUMO (Life Technologies). Expression in *E. coli* BL21-DE3 was induced with 0.2 mM isopropyl-β-D-thiogalactoside at 37˚C for 20 h. Cells were lysed with a micro-fluidizer (Microfluidics, Westwood, MA) and TtAgo isolated by HisTrap HP (GE Healthcare) chromatography. The amino-terminal six-histidine tag was cleaved from TtAgo using SUMO-protease (Life Technologies), and then the protein was further purified by HiTrap SP HP (GE Healthcare) chromatography. Finally, purified TtAgo was dialyzed into storage buffer (18 mM HEPES-KOH, pH 7.4, 100 mM potassium acetate, 3 mM magnesium acetate, 0.1 mM EDTA, 0.01% (w/v) Igepal CA-630, 5 mM dithiothreitol, 20% (w/v) glycerol). RISC was assembled by incubating

3 μM 16 nt, synthetic, single-stranded DNA oligonucleotide corresponding to the first 16 nt of let-7a with or without a 3′ Alexa555 dye with 1 μM TtAgo for 30 min at 75˚C in a reaction comprising storage buffer containing 0.01% (w/v; f.c.) Igepal CA-630 and 8% (w/v; f.c.) glycerol. Unassembled DNA guide was removed by passing the reaction through a Q Sepharose Fast Flow (GE Healthcare) spin column. Protein concentration was measured by Bradford assay. The concentration of active TtAgo was determined by stoichiometric binding as described (Wee et al., 2012).

Microscope Slide Preparation

Because the traditional polyethylene glycol (PEG) coating (Joo and Ha, 2012) resulted in high non-specific binding of RISC to the slide, we developed an alternative strategy using poly-L-lysine-graft-PEG copolymer (PLL-g-PEG) and heparin. PLL-g-PEG copolymer was prepared by dissolving 13 mg of poly-L-lysine (Sigma P7890) in 260 μL 0.1 M sodium bicarbonate to yield a ~50 $mg \cdot mL^{-1}$ solution. The solution was transferred to a vial containing 2 mg biotin-PEG-succinimidyl valerate MW 3,400 (Laysan Bio, Arab, AL) and vortexed to dissolve. This solution was then transferred to a vial containing 80 mg mPEG-succinimidyl valerate MW 2,000 (Laysan Bio) and vortexed to dissolve. The mixture was incubated at room temperature for 4 h with slow agitation, dialyzed against 2 L phosphate buffered saline buffer (10 mM disodium phosphate, 1.8 mM monopotassium phosphate,

137 mM sodium chloride, 2.7 mM potassium chloride, pH 7.4) for 12–16 h at 4°C, and then against 2 L deionized water for 12 h at 4°C using a 6–8 kDa MW cutoff membrane (Gerard Biotech, Oxford, OH). The solution (≤ 200 µL) was aliquoted into pre-weighed 1.5 mL microcentrifuge tubes. Aliquots were frozen in dry ice for 10 min, then dried in a SpeedVac rotary evaporator without heat for ≥ 2 h or until completely dry. Each tube was weighed again to determine the mass of dried PLL-g-PEG, which was then stored at –20°C.

Microscope slides (Gold Seal 24 × 60 mm, No. 1.5, Cat. #3423), and coverslips (Gold Seal 25 × 25 mm, No. 1, Cat. #3307) were cleaned by sonicating for 30 min in NanoStrip (KMG Chemicals, Houston, TX), followed by washing with at least 10 changes of deionized water. Clean slides and coverslips were stored in deionized water.

Slides and coverslips were dried with a stream of nitrogen. Using a plastic syringe tipped with a 200 µL pipette tip and filled with high vacuum grease (Dow Corning, Midland, MI), six ~1 mm diameter beads of vacuum grease were applied to the slide to create five parallel flow cells. The coverslip was placed on top of the slide and gently pressed down to ensure a good seal between the glass and vacuum grease and generate a ~0.3 mm gap between the slide and coverslip. Each flow cell volume was ~30 µL. Flow cells were filled with 2 $mg \cdot mL^{-1}$ PLL-g-PEG in 10 mM HEPES-KOH, pH 7.4, and incubated at room temperature for ≥1 h. Fresh slides were prepared for each day of imaging.

A syringe pump (KD Scientific, Holliston, MA) run in the withdrawal mode at 0.15 mL·min^{-1} was used to apply suction to the flow cell outlet to perform washes and flow in reagents. Immediately before each experiment, a flow cell was filled with 10 μg·mL^{-1} streptavidin (New England Biolabs), incubated for 2 min and washed three times with 60 μL LSE buffer (30 mM HEPES-KOH (pH 7.9), 120 mM potassium acetate, 3.5 mM magnesium acetate, 20% (w/v) glycerol, 1 mM dithiothreitol), and then incubated with 60 μL LSE buffer supplemented with 50 μg·mL^{-1} heparin (Sigma H4784) for 2 min before tethering target molecules to the slide surface.

Single-Molecule Microscopy

Imaging was performed on an IX81-ZDC2 zero-drift inverted microscope equipped with a cell^TIRF motorized multicolor TIRF illuminator with 491, 561, and 640 nm 100 W lasers and a 100×, 1.49 numerical aperture UAPON 100×OTIRF objective (Olympus, Tokyo, Japan). A digitally-controlled objective heater (Bioscience Tools, San Diego, CA) maintained objective temperature at 37˚C. Temperature on the slide surface was independently monitored by a small gauge thermocouple (Type E, 0.25 mm O.D., Omega Engineering Inc., Sutton, MA) inserted between the slide and coverslip. Images were recorded with two EM-CCD cameras (ImagEM, Hamamatsu Photonics, Hamamatsu, Japan). A dichroic image splitter (DC2, Photometrics, Tucson, AZ) was used to separate fluorescent emissions from spectrally distinct fluorescent dyes. All illumination

and acquisition parameters were controlled with Metamorph software (Molecular Devices, Sunnyvale, CA). The TIRF imaging system was isolated from floor vibrations with a Micro-g laboratory table (Technical Manufacturing Corporation, Peabody, MA).

The enzymatic oxygen scavenging system comprised 5 mM protocatechuic acid (PCA, Aldrich 37580) and 1 $U \cdot mL^{-1}$ *Pseudomonas sp.* protocatechuate 3,4-Dioxygenase (PCD, Sigma P8279; Crawford et al., 2008; Aitken et al., 2008). PCA was purified before use by two rounds of recrystallization from water. Triplet quenchers, trolox (Aldrich 238813), propyl gallate (Sigma P3130), and 4-nitrobenzyl alcohol (Aldrich N12821) were each added to 1 mM f.c. (Dave et al., 2009). Target molecules were tethered to the surface at a density of ~0.1 molecule/μm^2, ensuring that most molecules were individual, well isolated spots and minimizing clusters.

After a flow cell was treated with streptavidin and heparin, it was filled with ~50 pM target in LSE supplemented with 50 $\mu g \cdot mL^{-1}$ heparin, oxygen scavenging system and triplet quenchers. Target deposition was monitored by taking a series of images; once the desired density was achieved, the flow cell was washed twice with 50 µL LSE supplemented with oxygen scavenging system and triplet quenchers. To image RISC binding, continuous acquisition of frames was started simultaneously with flowing in RISC solution in LSE buffer

supplemented with oxygen scavenging system and triplet quenchers. Typically 2,000–4,000 frames were collected at 2–10 frames·s^{-1}.

Image Analysis

Image analysis was performed in MATLAB using custom scripts and the MATLAB co-localization analysis package (Friedman and Gelles, 2015). Images were recorded as uncompressed TIFF files and later merged into stacked TIFF files. Lateral drift of the surface was determined for each frame using target molecules as fiducial markers. Locations of target molecules were picked in the first frame acquired by a fast bandpass filter-based search algorithm (Crocker and Grier, 1996). Chosen target locations were visually examined, and locations not corresponding to individual target molecules removed. Then, the positions of target molecules were determined with higher accuracy by 2D Gaussian fitting. The positions of targets were then mapped onto the RISC channel as previously described (Friedman et al., 2013). Maps for this procedure were created using images of fluorescent streptavidin-labeled microspheres that were tethered to the biotinylated slide surface and that emit in both the target and RISC channels (40 nm diameter; Life Technologies F-8780). Coordinates of target locations in each frame in both the target and RISC channel were corrected for lateral drift. Detection of fluorescent spots and co-localization analysis in all frames was carried out as described (Friedman et al., 2013), with the criterion for target and RISC spot co-localization being within < 1 camera

pixel (160 nm) of the location of each other. This analysis yielded time intervals for RISC binding events to each target molecule. The same analysis was carried out for “dark” locations on the slide surface that did not contain target molecules; these served as a control for non-specific binding of RISC to the slide surface. Typically, 400–700 target molecule locations and an equal number of “dark” locations were analyzed.

Correction for Non-Specific Events

In TIRF-based, single-molecule co-localization experiments, one of the interacting molecules must be tethered to a slide surface for its stable association with a freely diffusing interaction partner to be visualized as the co-localization of their spectrally distinct fluorescent emissions. Despite efforts to minimize non-specific binding by coating the slide surface with hydrophilic polymers (e.g., polyethylene glycol), a background of non-specific binding is always present, and may be high depending on the nature and concentration of fluorescent molecules. Both on-rate and off-rate measurements are affected by non-specific binding and need to be corrected for the non-specific background. In addition, off-rate measurements need to be corrected for photobleaching, because RISC departure is indistinguishable from photobleaching.

On-Rate Correction for Non-Specific Arrivals

To subtract non-specific arrivals from observed arrivals (a mixture of specific and non-specific arrivals), the specific arrivals may be represented in statistical

terms as the intersection of two events for a given time bin (e.g., each acquired frame): A, a specific arrival occurs, and B, a non-specific arrival has never occurred in any of the prior time bins. Probability of event A for time interval between t and $(t + \Delta t)$, $P_A(t< T< (t + \Delta t))$, is given by $f_S(t) \cdot \Delta t$, where $f_S(t)$ is the probability density function (PDF), and Δt is a small time interval. For a single exponential process, $f_S(t) = k_S \cdot \exp(-k_S \cdot t)$, where k_S is the rate constant for specific arrivals. The probability of the event B, $P_B(T > t)$, is given by the survival function $S_{NS}(t) = 1 - F_{NS}(t)$, where $F_{NS}(t)$ is the cumulative distribution function (CDF). Assuming a single exponential process for non-specific arrivals, $F_{NS}(t) = 1 - \exp(-k_{NS} \cdot t)$. $F_{NS}(t)$, and therefore $S_{NS}(t)$ can be obtained from the control experiments where “dark” locations that do not contain target molecules are selected, and so all RISC arrivals must be non-specific.

Conversely, non-specific arrivals can be represented as the intersection of events A′ and B′, where A′ is the event where a non-specific arrival occurs in the time interval $t < T < (t + \Delta t)$, and B′ is the event where a specific arrival has never occurred in any of the prior intervals. Assuming that non-specific arrivals follow exponential distribution, probability of A′ is given by $\Delta t \cdot f_{NS}(t) = \Delta t \cdot k_{NS} \cdot \exp(-k_{NS} \cdot t)$, and probability of B′ is given by $S_S(t) = \exp(-k_S \cdot t)$, where $f_{NS}(t)$ is the PDF for non-specific arrivals, and $S_S(t)$ is the survival function for specific arrivals, and k_{NS} is the rate constant for non-specific arrivals.

Since A and B are independent events, their joint probability is the product of their probabilities:

$$\mathrm{P(AUB)} = \mathrm{P(A)} \cdot \mathrm{P(B)} = \Delta \mathrm{t} \cdot k_S \cdot \exp\left(-\mathrm{k_S} \cdot \mathrm{t}\right) \cdot \exp\left(-k_{NS} \cdot \mathrm{t}\right))$$

$$= \Delta \mathrm{t} \cdot k_S \cdot \exp\left(-(k_S + k_{NS}) \cdot \mathrm{t}\right)$$

The same is true for A′ and B′, and

$$\mathrm{P(A'\ UB'\ )} = \mathrm{P(A'\ )} \cdot \mathrm{P(B'\ )} = \Delta \mathrm{t} \cdot k_{NS} \cdot \exp\left(-k_{NS} \cdot \mathrm{t}\right) \cdot \exp\left(-k_S \cdot \mathrm{t}\right))$$

$$= \Delta \mathrm{t} \cdot k_{NS} \cdot \exp\left(-(k_S + k_{NS}) \cdot \mathrm{t}\right)$$

In our experiments to measure on-rate, only the first arrival at each target location is counted, and this arrival can be either specific, or non-specific. Therefore, the probability to count any arrival in a given time interval is the union of A∪B and A′∪B′ and is given by

$$P_{exp} = \mathrm{P}\big((\mathrm{AUB})\mathrm{U}\ (\mathrm{A'\ UB'\ })\big) = \mathrm{P(A)} \cdot \mathrm{P(B)} + \mathrm{P(A'\ )} \cdot \mathrm{P(B'\ )} = (k_S + k_{NS}) \cdot \Delta \mathrm{t} \cdot \exp(-(k_S + k_{NS}) \cdot \mathrm{t})$$

Therefore the corresponding PDF for experimental (mixed specific and non-specific) arrivals is

$$\mathrm{f_{exp}(t)} = (\mathrm{k_S} + k_{NS}) \cdot \exp(-(k_S + k_{NS}) \cdot \mathrm{t}),$$

and the CDF is

$$\mathrm{F_{exp}(t)} = 1 - \exp(-(k_S + k_{NS}) \cdot \mathrm{t}) \quad (1)$$

Thus, the experimentally measured rate is just a sum of specific and non-specific arrival rates. Therefore the experimental PDF for mixed arrivals can be expressed as:

$$\mathrm{f_{exp}(t)} \;=\; ((k_S + k_{NS})/\, k_S \cdot k_{NS}) \cdot \mathrm{f_S(t)} \cdot \mathrm{f_{NS}(t)},$$

which turns out not to be a sum but rather a weighted product of specific and non-specific PDFs.

In many cases, we observe that only a fraction of “dark” locations on the slide surface that are picked to measure non-specific rates will ever receive non-specific arrivals, apparently due to non-uniform coating of the slide surface with PLL-g-PEG. This fraction (h, for ‘hot’ locations) can be easily determined from the amplitude of the cumulative binding curve for non-specific arrivals. Using the same reasoning as before, we can show that in this case the experimental CDF for mixed arrivals is given by

$$\mathrm{F_{exp}(t)} \;=\; 1 - (1 - \mathrm{h}) \cdot \exp(-k_S \cdot \mathrm{t}) \;-\; \mathrm{h} \cdot \exp(-(k_S + k_{NS}) \cdot \mathrm{t}) \qquad (2)$$

Each on-rate measurement was analyzed by picking two sets of locations: first, specific locations where a target molecule was seen, and second, non-specific locations that did not contain target molecules. The waiting times before first RISC arrival in all locations were measured and converted into cumulative form C(t), where for each waiting time t, arrivals that occurred before t were summed to make up the cumulative number of arrivals. For the non-specific arrivals, the

rate (k_{NS}), and where applicable, the fraction of locations that exhibited non-specific arrivals (h) was determined from single-exponential fitting:

$$\mathrm{C_{NS}(t)} = \mathrm{h} \cdot \mathrm{N_{NS}} \cdot (1 - \exp(-k_{NS} \cdot \mathrm{t})),$$

where $\mathrm{N_{NS}}$ is the number of non-specific locations picked for analysis.

Next, the specific rate for RISC arrivals to the target-specific locations was determined by curve fitting using the known values of h and k_{NS}:

$$\mathrm{C_S(t)} = \mathrm{N_S} \cdot \mathrm{F_{exp}(t)},$$

where $\mathrm{N_S}$ is the number of target-specific locations that bind RISC (determined from fitting), and $\mathrm{F_{exp}(t)}$ is given by (2).

Off-Rate Correction for Non-Specific Events

Some of the observed RISC binding events are false and come from non-specific binding to the surface fortuitously close enough to the target location to be counted as specific binding according to our co-localization criteria. The apparent cumulative counts of RISC dissociation events are the sum of counts from target-specific RISC events and non-specific events:

$$C_{\exp(t)} = C_{RISC}(t) + C_{NS}(t)$$

where $C_{NS}(t)$ corresponds to counts for RISC molecules bound non-specifically in target locations, and their contribution can be determined using “dark” locations. We find that these non-specific events display double-exponential kinetics, and:

$$C_{NS}(t) = A(1 - \exp(k_{NS1} \cdot t)) + B(1 - \exp(k_{NS2} \cdot t)),$$

where k_{NS1} and k_{NS1} are rates for non-specifically bound RISC, and A and B are their respective amplitudes. We find that RISC dissociation kinetics corrected for non-specific binding follows a single exponential, and the expression for fitting experimentally observed cumulative dissociation counts is given by

$$C_{exp}(t) = N_{RISC} \cdot (1 - \exp(-k_{RISC} \cdot t)) + A(1 - \exp(k_{NS1} \cdot t)) + B(1 - \exp(k_{NS2} \cdot t)),$$

where N_{RISC} is amplitude for specific RISC events, k_{RISC} is their rate, and the rest of parameters are determined from the fitting of data for "dark" locations.

Off-Rate Correction for Photobleaching

Equation (1), which was used to correct the on-rates for non-specific RISC binding, is also applicable to correcting RISC dwell times for photobleaching, since binding and photobleaching are independent processes and have characteristic exponential rates. Therefore, the observed cumulative rate for the apparent dwell times of RISC is given by

$$\mathrm{F_{obs}(t)} = 1 - \exp\left(-\left(k_{photobleaching} + k_{RISC}\right) \cdot \mathrm{t}\right), \quad (3)$$

where k_{RISC} is the dissociation rate of target-bound RISC, corrected for non-specific events as described above, and $k_{photobleaching}$ is the rate of photobleaching. The photobleaching rate under our standard imaging conditions was determined in two ways. First, because the dwell times of RISC bound to the seed-matched target (Figure S3.5C) was of the order of hundreds of

seconds, the apparent dwell times were clearly limited by photobleaching. From this data, we determined a photobleaching rate of ~0.06 s^{-1} under our standard acquisition conditions. Second, the same analysis was carried out for the target that was fully complementary to the guide in RISC using those molecules that were bound by RISC but remained uncleaved in the experiments shown in Figure 3.5B; RISC molecules bound to these targets were eventually photobleached. Both approaches produced the same photobleaching rate. We conclude that our photobleaching rate was reproducible between experiments, and that $k_{photobleaching} \approx 0.06\ s^{-1}$. Experiments performed to obtain k_{off} for RISC bound to 3′ product mimics, seed + t9, and seed + t9t10 targets were performed under low time-resolution conditions (1.5 frames·s^{-1}) where our photobleaching rate was determined to be ~0.005 s^{-1}. To obtain the corrected dissociation rate for RISC, the known photobleaching rate was subtracted from the observed dissociation rate for RISC.

Nearest Neighbor Analysis

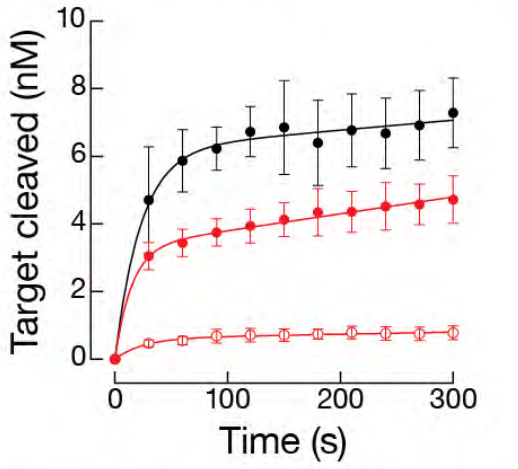
$\Delta G_{37℃} = -RT \ln(1/K_D)$, where R = 1.987 cal·$K^{-1}$·$mol^{-1}$ and T = 310.15 K. Theoretical $\Delta G_{37℃}$ was calculated from nearest neighbor values using RNAstructure 5.6 (Reuter and Mathews, 2010; Turner and Mathews, 2010).

ACKNOWLEDGEMENTS

We thank members of the Zamore and Moore laboratories for helpful discussions and comments on the manuscript; Jeff Gelles and Larry Friedman for experimental and analytical advice; and Johnson Chung for scripts to create overlay movies (available online, http://dx.doi.org/10.1016/j.cell.2015.06.029). This work was supported in part by National Institutes of Health grants GM053007 and GM053007 to M.J.M. and GM62862 and GM65236 to P.D.Z.

Figure S3.1

A



Purified Mouse let-7a AGO2-RISC

Preparation	Active RISC (nM)
3′ OH	5.5 ± 0.7
3′ Alexa555	3.5 ± 0.4
3′ Alexa555	0.67 ± 0.16

$$F(t) = E \times \frac{a^2}{(a+b)^2}\left(1 - e^{-(a+b)t}\right) + E \times \frac{ab}{(a+b)}t$$

B

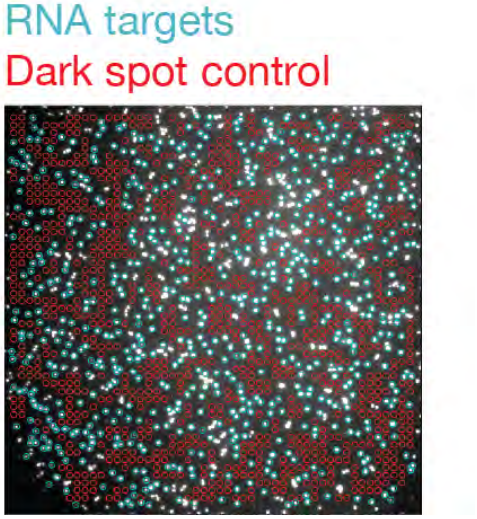


Figure S3.1: Controls and Concentration Determination for mouse AGO2

(A) Pre-steady state kinetics was used to estimate the active concentration of let-7a-RISC. The substrate concentration used for unlabeled RISC was 200 nM (*n* = 3). For both RISC preparation 1 (*n* = 5) and preparation 2 (*n* = 3), 100 nM substrate was used in the assay. Values are mean ± S.D.; the equation used to fit each curve is shown below the table. (B) Image representing the selection (blue squares around molecules) of the RNA targets used to analyze RISC activity. These selections are referred to as areas of interest (AOIs).

Figure S3.2

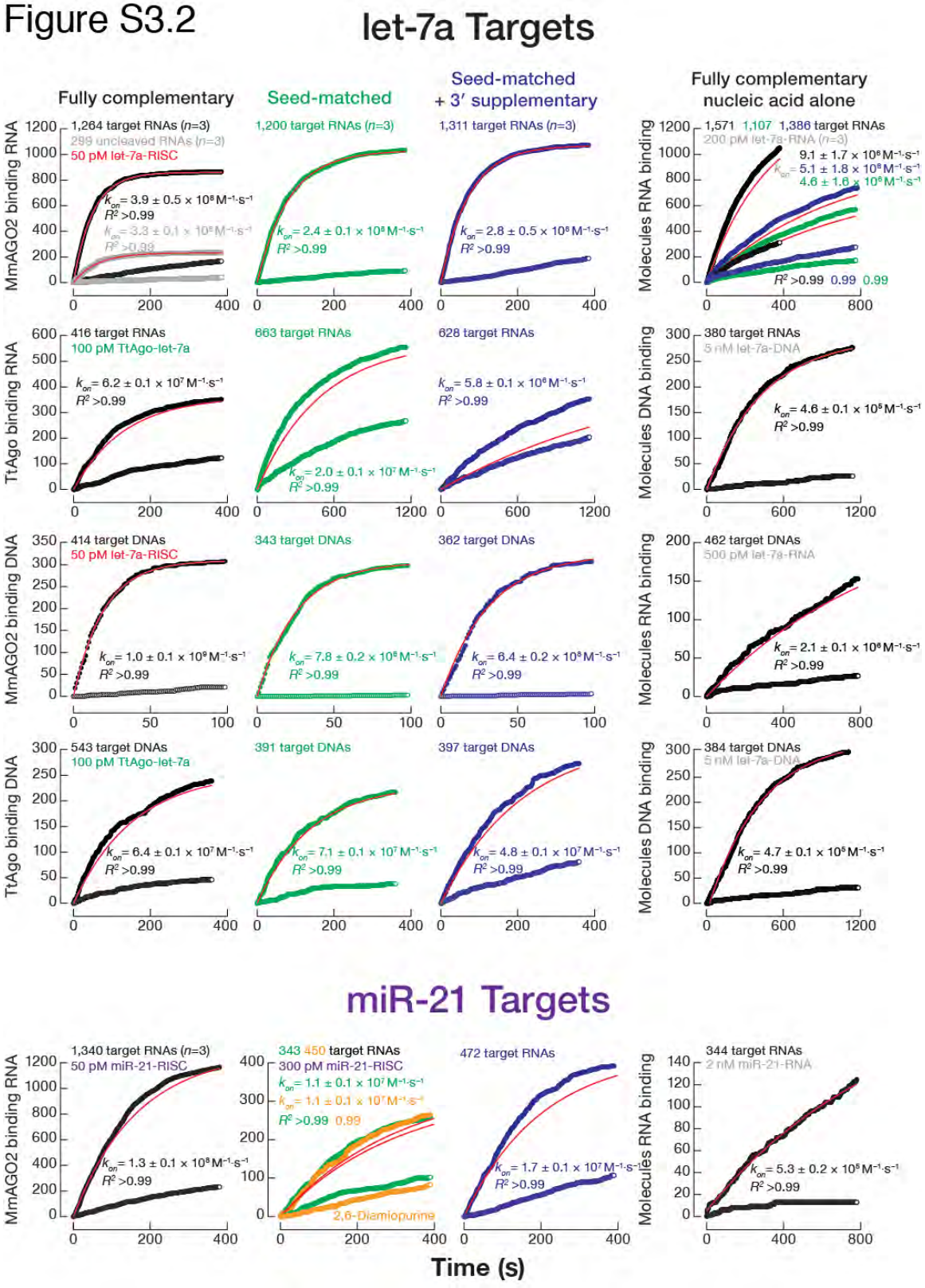

let-7a Targets
Fully complementary
Seed-matched
Seed-matched + 3′ supplementary
Fully complementary nucleic acid alone
MmAGO2 binding RNA
1,264 target RNAs (n=3)
299 uncleaved RNAs (n=3)
50 pM let-7a-RISC
k_on = 3.9 ± 0.5 × 10^8 M^-1·s^-1
R^2 >0.99
k_on = 3.3 ± 0.1 × 10^8 M^-1·s^-1
R^2 >0.99
1,200 target RNAs (n=3)
k_on = 2.4 ± 0.1 × 10^8 M^-1·s^-1
R^2 >0.99
1,311 target RNAs (n=3)
k_on = 2.8 ± 0.5 × 10^8 M^-1·s^-1
R^2 >0.99
Molecules RNA binding
1,571 1,107 1,386 target RNAs
200 pM let-7a-RNA (n=3)
9.1 ± 1.7 × 10^8 M^-1·s^-1
k_on = 5.1 ± 1.8 × 10^8 M^-1·s^-1
4.6 ± 1.8 × 10^8 M^-1·s^-1
R^2 >0.99 0.99 0.99
TtAgo binding RNA
416 target RNAs
100 pM TtAgo-let-7a
k_on = 6.2 ± 0.1 × 10^7 M^-1·s^-1
R^2 >0.99
663 target RNAs
k_on = 2.0 ± 0.1 × 10^7 M^-1·s^-1
R^2 >0.99
628 target RNAs
k_on = 5.8 ± 0.1 × 10^6 M^-1·s^-1
R^2 >0.99
Molecules DNA binding
380 target RNAs
5 nM let-7a-DNA
k_on = 4.6 ± 0.1 × 10^5 M^-1·s^-1
R^2 >0.99
MmAGO2 binding DNA
414 target DNAs
50 pM let-7a-RISC
k_on = 1.0 ± 0.1 × 10^9 M^-1·s^-1
R^2 >0.99
343 target DNAs
k_on = 7.8 ± 0.2 × 10^8 M^-1·s^-1
R^2 >0.99
362 target DNAs
k_on = 6.4 ± 0.2 × 10^8 M^-1·s^-1
R^2 >0.99
Molecules RNA binding
462 target DNAs
500 pM let-7a-RNA
k_on = 2.1 ± 0.1 × 10^8 M^-1·s^-1
R^2 >0.99
TtAgo binding DNA
543 target DNAs
100 pM TtAgo-let-7a
k_on = 6.4 ± 0.1 × 10^7 M^-1·s^-1
R^2 >0.99
391 target DNAs
k_on = 7.1 ± 0.1 × 10^7 M^-1·s^-1
R^2 >0.99
397 target DNAs
k_on = 4.8 ± 0.1 × 10^7 M^-1·s^-1
R^2 >0.99
Molecules DNA binding
384 target DNAs
5 nM let-7a-DNA
k_on = 4.7 ± 0.1 × 10^6 M^-1·s^-1
R^2 >0.99
miR-21 Targets
MmAGO2 binding RNA
1,340 target RNAs (n=3)
50 pM miR-21-RISC
k_on = 1.3 ± 0.1 × 10^8 M^-1·s^-1
R^2 >0.99
343 450 target RNAs
300 pM miR-21-RISC
k_on = 1.1 ± 0.1 × 10^7 M^-1·s^-1
k_on = 1.1 ± 0.1 × 10^7 M^-1·s^-1
R^2 >0.99 0.99
2,6-Diaminopurine
472 target RNAs
k_on = 1.7 ± 0.1 × 10^7 M^-1·s^-1
R^2 >0.99
Molecules RNA binding
344 target RNAs
2 nM miR-21-RNA
k_on = 5.3 ± 0.2 × 10^5 M^-1·s^-1
R^2 >0.99
Time (s)

Figure S3.2: Argonaute Binds to Its Target Faster than Guide Strand Alone

Binding curves for mouse AGO2 or TtAgo compared to guide strand alone. Solid markers: the cumulative number of molecules binding for the first time to a single RNA or DNA target. Open markers: the cumulative number of molecules binding to regions of slide surface that did not contain an RNA or DNA target. The curve in red shows the rate of binding after correcting for non-specific binding.

Figure S3.3

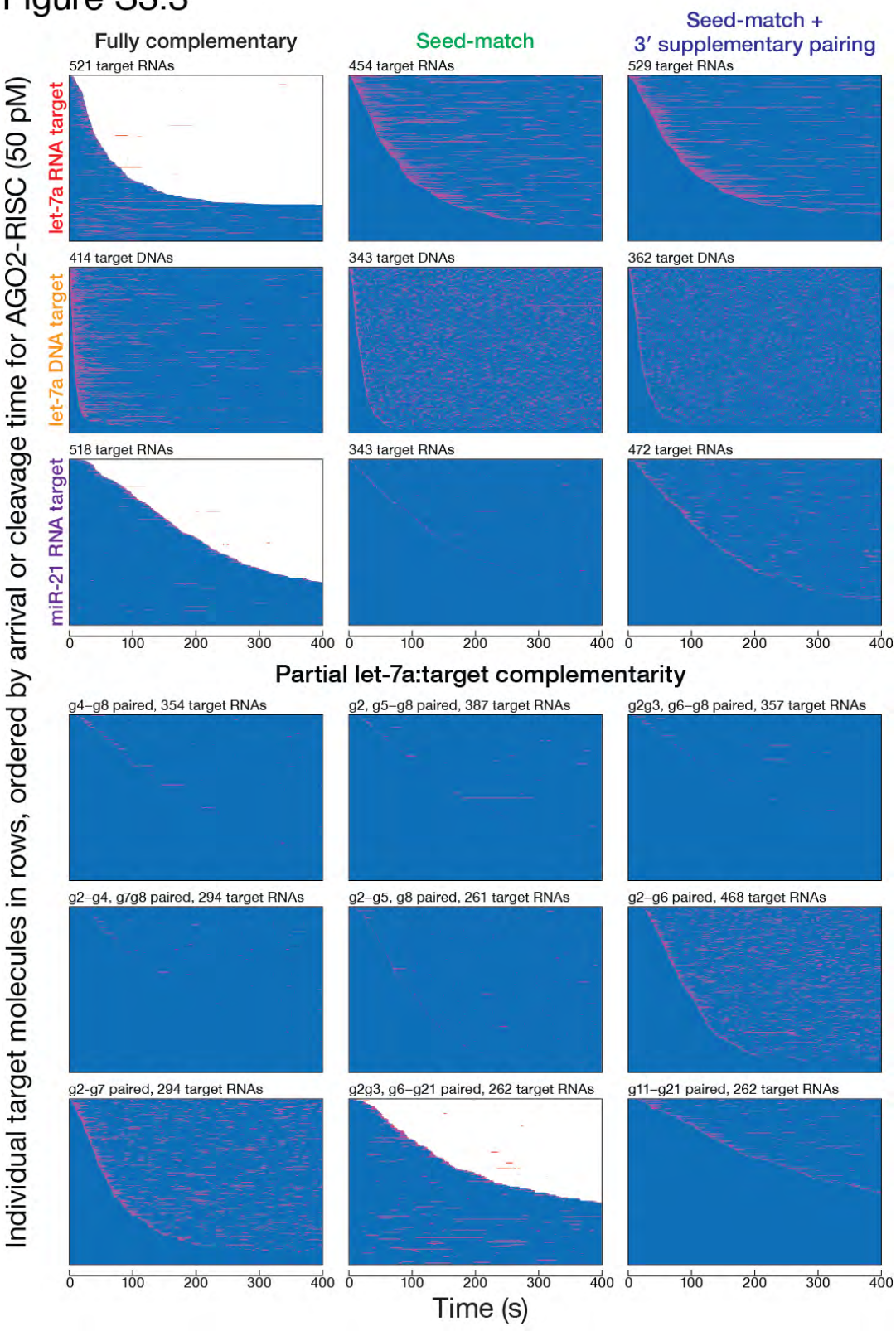

Fully complementary
Seed-match
Seed-match +
3′ supplementary pairing
521 target RNAs
454 target RNAs
529 target RNAs
let-7a RNA target
414 target DNAs
343 target DNAs
362 target DNAs
let-7a DNA target
518 target RNAs
343 target RNAs
472 target RNAs
miR-21 RNA target
0
100
200
300
400
Partial let-7a:target complementarity
g4–g8 paired, 354 target RNAs
g2, g5–g8 paired, 387 target RNAs
g2g3, g6–g8 paired, 357 target RNAs
g2–g4, g7g8 paired, 294 target RNAs
g2–g5, g8 paired, 261 target RNAs
g2–g6 paired, 468 target RNAs
g2-g7 paired, 294 target RNAs
g2g3, g6–g21 paired, 262 target RNAs
g11–g21 paired, 262 target RNAs
Individual target molecules in rows, ordered by arrival or cleavage time for AGO2-RISC (50 pM)
Time (s)

Figure S3.3: Rastergrams for Mouse AGO2 to RNA or DNA Targets

Rastergram summary of traces of individual targets molecules for different guide:target pairings for mouse AGO2. A representative rastergram is shown for each of the experiments performed with canonical let-7a RNA targets and miR-21 perfect RNA target.

Figure S3.4

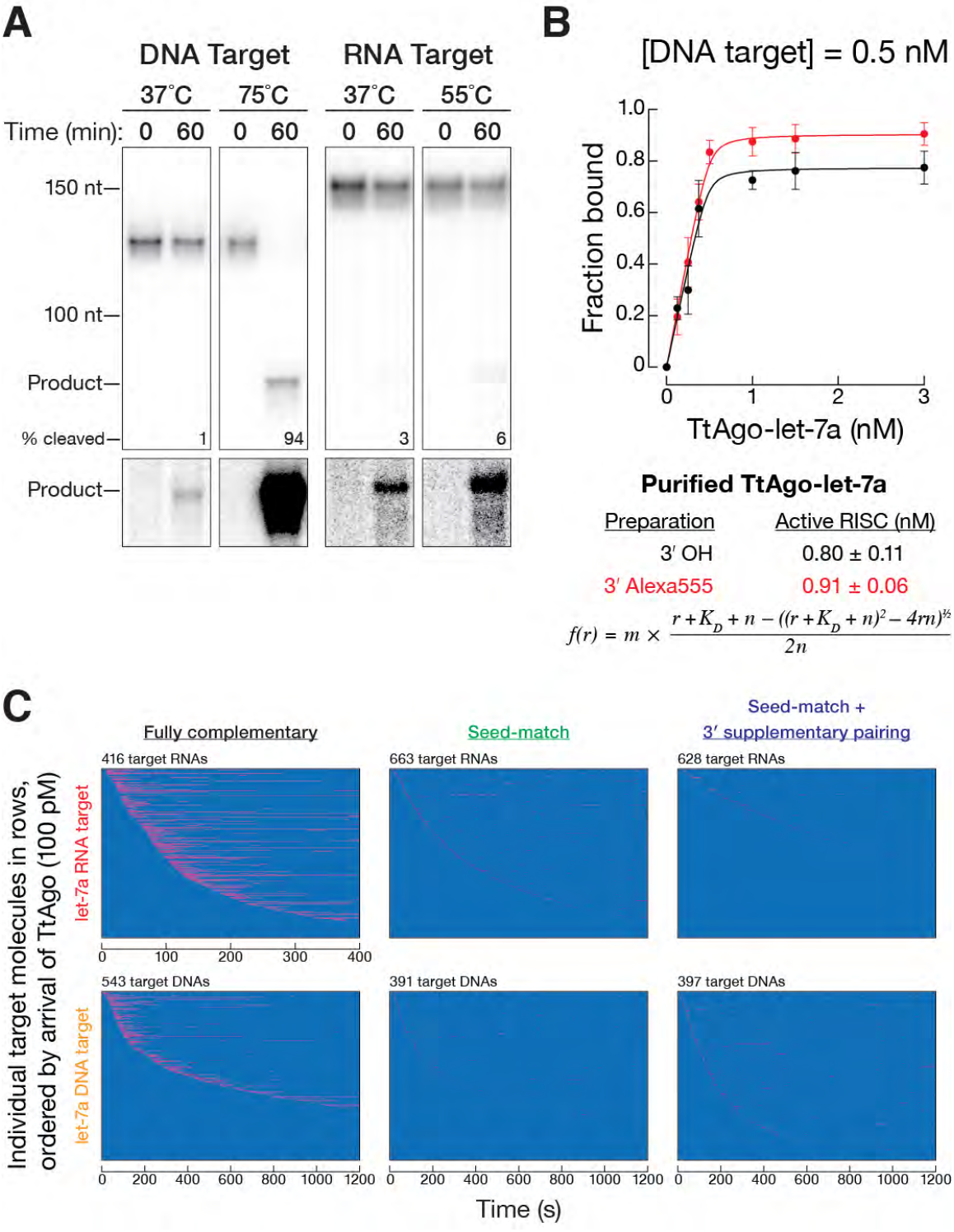

A
DNA Target
RNA Target
37˚C
75˚C
37˚C
55˚C
Time (min): 0 60 0 60 0 60 0 60
150 nt—
100 nt—
Product—
% cleaved—
1
94
3
6
Product—
B
[DNA target] = 0.5 nM
Fraction bound
1.0
0.8
0.6
0.4
0.2
0
0
1
2
3
TtAgo-let-7a (nM)
Purified TtAgo-let-7a
Preparation
Active RISC (nM)
3′ OH
0.80 ± 0.11
3′ Alexa555
0.91 ± 0.06
f(r) = m × (r + K_D + n − ((r + K_D + n)^2 − 4rn)^½) / 2n
C
Fully complementary
Seed-match
Seed-match +
3′ supplementary pairing
416 target RNAs
663 target RNAs
628 target RNAs
let-7a RNA target
543 target DNAs
391 target DNAs
397 target DNAs
let-7a DNA target
Individual target molecules in rows, ordered by arrival of TtAgo (100 pM)
0 100 200 300 400
0 200 400 600 800 1000 1200
Time (s)

Figure S3.4: Activity, Concentration determination, and Rastergrams for *Thermus thermophilus* Argonaute for DNA and RNA targets

(A) Target cleavage by of TtAgo at 37˚C, 55˚C, and 75˚C. A 5′ ^{32}P-radiolabeled synthetic DNA or an in vitro transcribed RNA was incubated with purified let-7a-guided TtAgo-RISC. TtAgo cleavage activity after 1 h at 37˚C was ~2 nM for either RNA or DNA, which is less than a single turnover (~ 35 nM let-7a-TtAgo was used). A digital over-exposure is shown below the linear exposure of the gel. (B) Stoichiometric target DNA binding analysis of TtAgo RISC guided with a 16 nt DNA corresponding to the first 16 nt of let-7a. (C) Rastergram summary of traces of individual target molecules for different guide:target pairings for *Thermus thermophilus* Argonaute bound to a 16-nt DNA guide.

Figure S3.5

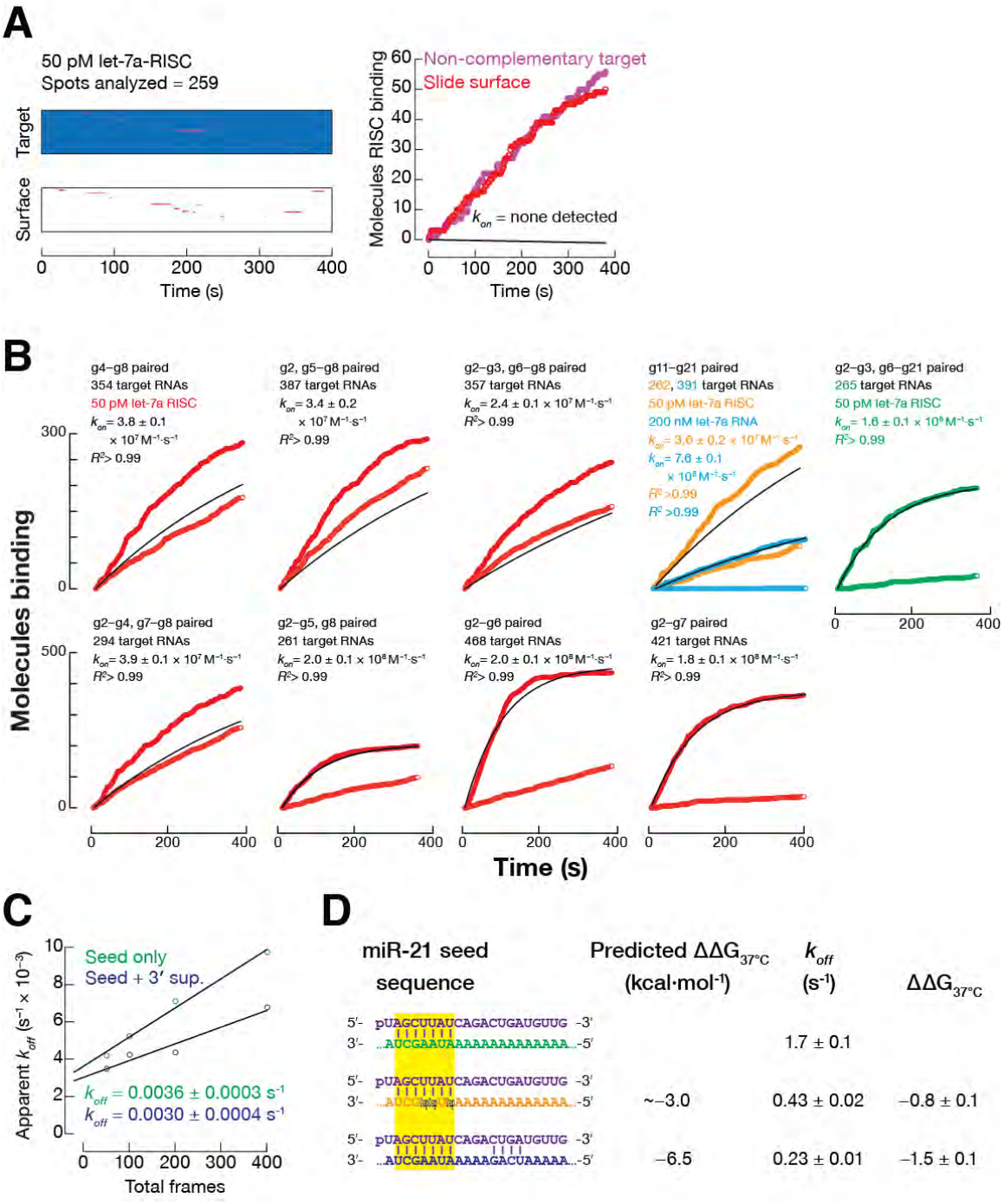

A
50 pM let-7a-RISC
Spots analyzed = 259
Target
Surface
0 100 200 300 400
Time (s)
Molecules RISC binding
Non-complementary target
Slide surface
k_{on} = none detected
0 10 20 30 40 50 60
0 100 200 300 400
Time (s)
B
Molecules binding
g4–g8 paired
354 target RNAs
50 pM let-7a RISC
k_{on} = 3.8 ± 0.1 × 10⁷ M⁻¹·s⁻¹
R^2 > 0.99
g2, g5–g8 paired
387 target RNAs
k_{on} = 3.4 ± 0.2 × 10⁷ M⁻¹·s⁻¹
R^2 > 0.99
g2–g3, g6–g8 paired
357 target RNAs
k_{on} = 2.4 ± 0.1 × 10⁷ M⁻¹·s⁻¹
R^2 > 0.99
g11–g21 paired
262, 391 target RNAs
50 pM let-7a RISC
k_{on} = 3.6 ± 0.2 × 10⁷ M⁻¹·s⁻¹
200 nM let-7a RNA
k_{on} = 7.6 ± 0.1 × 10⁸ M⁻¹·s⁻¹
R^2 >0.99
R^2 >0.99
g2–g3, g6–g21 paired
265 target RNAs
50 pM let-7a RISC
k_{on} = 1.6 ± 0.1 × 10⁸ M⁻¹·s⁻¹
R^2 > 0.99
g2–g4, g7–g8 paired
294 target RNAs
k_{on} = 3.9 ± 0.1 × 10⁷ M⁻¹·s⁻¹
R^2 > 0.99
g2–g5, g8 paired
261 target RNAs
k_{on} = 2.0 ± 0.1 × 10⁸ M⁻¹·s⁻¹
R^2 > 0.99
g2–g6 paired
468 target RNAs
k_{on} = 2.0 ± 0.1 × 10⁸ M⁻¹·s⁻¹
R^2 > 0.99
g2–g7 paired
421 target RNAs
k_{on} = 1.8 ± 0.1 × 10⁸ M⁻¹·s⁻¹
R^2 > 0.99
0 200 400
Time (s)
C
Apparent k_{off} (s⁻¹ × 10⁻³)
Seed only
Seed + 3′ sup.
k_{off} = 0.0036 ± 0.0003 s⁻¹
k_{off} = 0.0030 ± 0.0004 s⁻¹
0 2 4 6 8 10
0 100 200 300 400
Total frames
D
miR-21 seed sequence
Predicted ΔΔG37°C (kcal·mol⁻¹)
k_{off} (s⁻¹)
ΔΔG37°C
5′- pUAGCUUAUCAGACUGAUGUUG -3′
3′- ...AUCGAAUAAAAAAAAAAAAA...-5′
1.7 ± 0.1
5′- pUAGCUUAUCAGACUGAUGUUG -3′
3′- ...AUCGAAUAAAAAAAAAAAAA...-5′
~–3.0
0.43 ± 0.02
–0.8 ± 0.1
5′- pUAGCUUAUCAGACUGAUGUUG -3′
3′- ...AUCGAAUAAAAGACUAAAAA...-5′
–6.5
0.23 ± 0.01
–1.5 ± 0.1

Figure S3.5: Mouse AGO2-RISC Binding to and Departing from Partially Paired or Seed-Mismatched Targets

(A) At left, rastergrams for let-7a-guided RISC binding to a non-complementary target (magenta on blue, with target) or the slide surface (red on white, with no target). At right, cumulative binding of let-7a-guided RISC to a non-complementary target. Curve in black is the fit to the cumulative distribution after subtracting non-specific binding. No binding was detected for an RNA target without complementarity to the let-7a guide strand. (B) Cumulative binding curves for targets with disrupted pairing to let-7a mouse RISC. Seed-matched targets bearing dinucleotide mismatches to the seed (red), target pairing t11-t21 (orange represents mouse AGO2-RISC and teal represents RNA alone), perfectly paired target with dinucleotide mismatch to g4g5 (green). (C) Determination of k_{off} for let-7a-RISC for seed-matched and seed-matched plus 3′ supplementary pairing targets. k_{off} for both targets was slower than the rate of photobleaching (k = 0.06 s^{-1}, τ ~17 s) in our standard acquisition conditions. To determine k_{off} for let-7a-RISC, the amount of laser exposure was varied by changing the frame rate. The y-intercept for apparent k_{off} versus laser exposure was taken to be k_{off}. (D) Measuring k_{off} for miR-21-RISC for seed-matched target, seed-matched target containing three 2,6 diaminopurine nucleotides (Dap) in place of adenosines, and seed-matched plus four 3′ complementary nucleotides targets. Predicted ΔΔG was determined by nearest neighbor analysis. k_{off} was corrected for the rate of photobleaching. Measured ΔΔG was calculated from the experimentally determined on- and off-rates: $\Delta\Delta G = \Delta G - \Delta G^{\text{seed-match}} = RT\ln(K_D / K_D^{\text{seed-match}})$ where $K_D = k_{off} / k_{on}$.

Figure S3.6

Guide	Target	Target Size	AGO2-RISC k_{on} ($M^{-1}\cdot s^{-1}$)
RNA	RNA	141 nt + 148 DNA	$3.9 \pm 0.5 \times 10^8$
RNA	RNA	28 nt	$2.3 \pm 0.1 \times 10^8$
RNA	RNA	141 nt + 148 DNA	$2.4 \pm 0.1 \times 10^8$
RNA	RNA	28 nt	$1.5 \pm 0.1 \times 10^8$

Figure S3.6: Comparison of Target length and k_{on} for let-7a AGO2-RISC.

The k_{on} for let-7a AGO2-RISC was measured under standard conditions at 37°C. The 28 nt RNA targets contained a single Alexa647 dye at the 3′ end and were attached to the slide via biotin streptavidin (Table S3.1). Target cleavage was not measured because the perfect target contained 2′-*O*-methyl at t10 and t11 and contained a phosphorothioate linkage between t10 and t11.

Figure S3.7

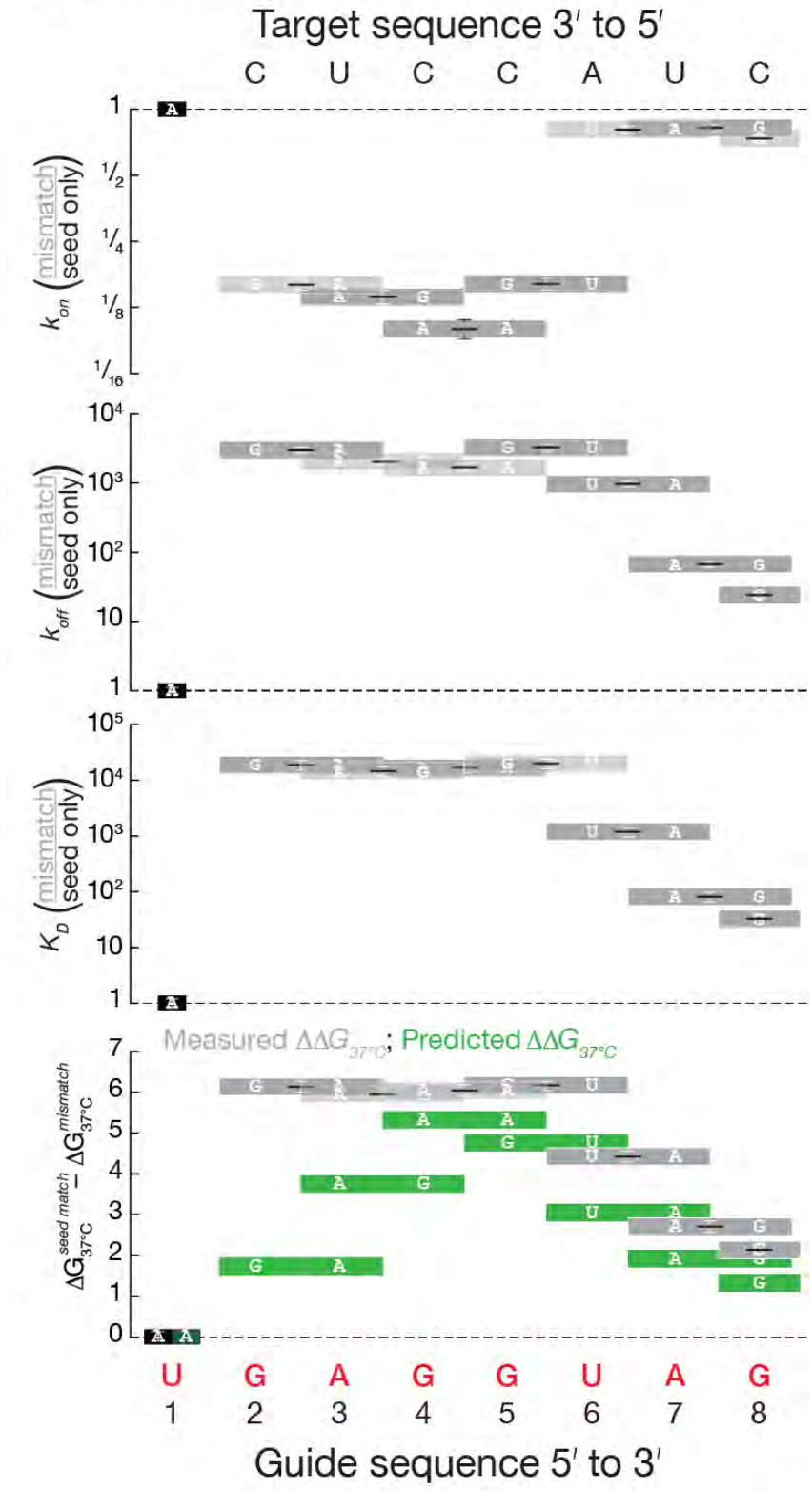

Target sequence 3′ to 5′
C U C C A U C
k_{on} (mismatch / seed only)
k_{off} (mismatch / seed only)
K_D (mismatch / seed only)
Measured ΔΔG37°C; Predicted ΔΔG37°C
$\Delta G^{seed\ match}_{37°C} - \Delta G^{mismatch}_{37°C}$
U G A G G U A G
1 2 3 4 5 6 7 8
Guide sequence 5′ to 3′

Figure S3.7: Target Mismatches to the Seed Decrease AGO2-RISC k_{on} and Increase k_{off}

Comparison of let-7a-guided AGO2-RISC binding to targets with dinucleotide mismatches to the seed sequence. The nucleotides present in the seed mismatched positions are indicated for each value. Error bars represent the propagated error determined from the comparison of seed mismatch over complete seed only pairing (black bar). Relative $\Delta\Delta G_{37°C}$ was predicted by nearest neighbor analysis (Materials and Methods).

Figure S3.8

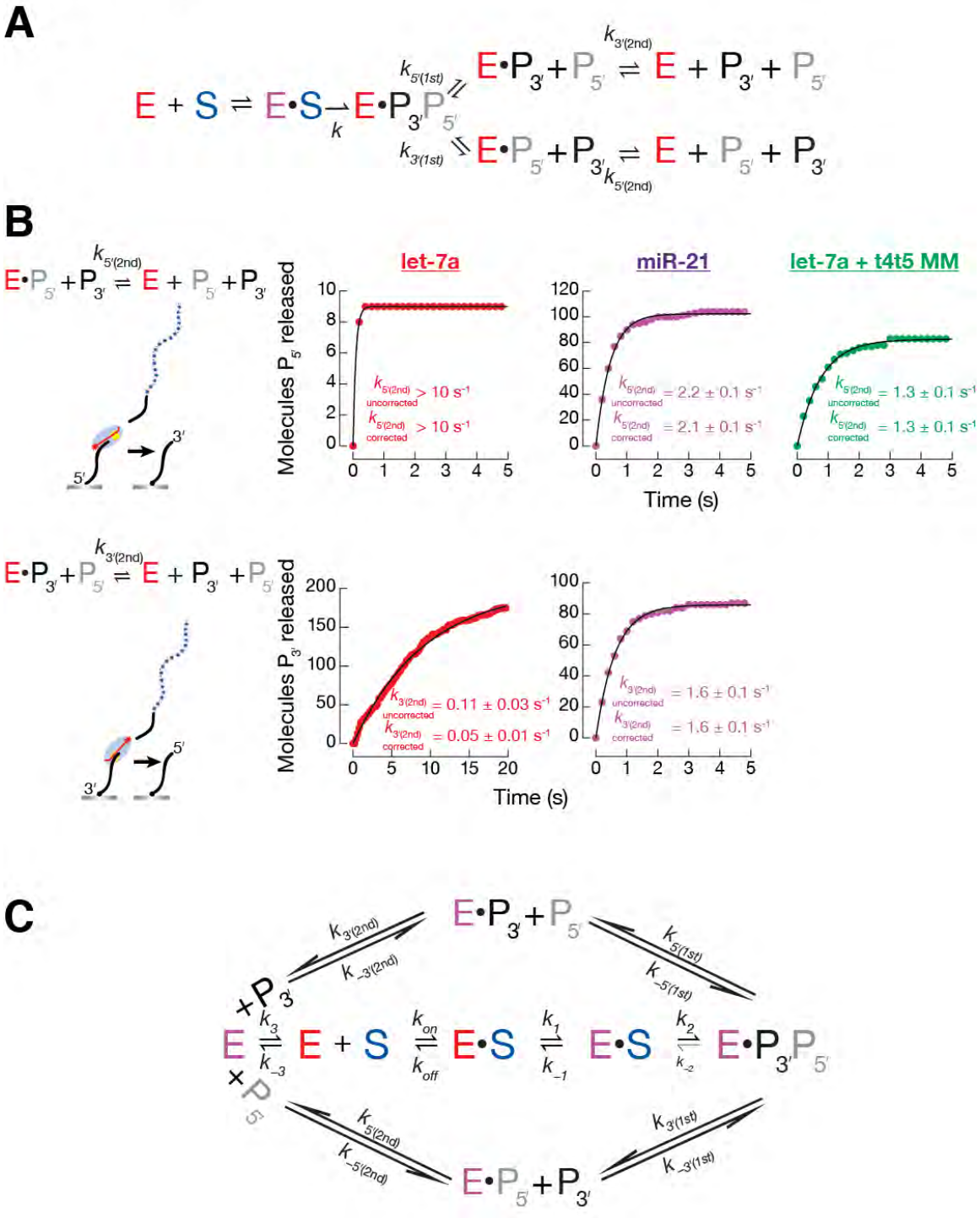

A
B
C
let-7a
miR-21
let-7a + t4t5 MM
Molecules P5′ released
Molecules P3′ released
Time (s)
k5′(2nd) > 10 s-1 uncorrected
k5′(2nd) > 10 s-1 corrected
k5′(2nd) = 2.2 ± 0.1 s-1 uncorrected
k5′(2nd) = 2.1 ± 0.1 s-1 corrected
k5′(2nd) = 1.3 ± 0.1 s-1 uncorrected
k5′(2nd) = 1.3 ± 0.1 s-1 corrected
k3′(2nd) = 0.11 ± 0.03 s-1 uncorrected
k3′(2nd) = 0.05 ± 0.01 s-1 corrected
k3′(2nd) = 1.6 ± 0.1 s-1 uncorrected
k3′(2nd) = 1.6 ± 0.1 s-1 corrected

Figure S3.8: Seed Base-Pairing Stability Determines Product Release

(A) Schematic of RNAi and the two product release steps. (B) Cumulative product release of 5′ or 3′ cleavage products after RNAi. (C) Schematic of RNAi reaction model. To account for additional steps that may explain *k* rate the addition of a conformation step that represents the two-state model (Tomari and Zamore, 2005) and chemistry step for endonucleolytic cleavage is expressed in the schematic while representing the two different branches of product release. “E” represents AGO2-RISC, when E is red it is in the first state, ready for target binding and when E is magenta it is in the second state conformation or cleavage competent conformation. “S” represents the substrate, uncleaved target and “P” represents cleavage products (subscript denotes 5′ or 3′). The *k* rate measured by global fit analysis is $k = k_1$ (conformation change) + k_2 (chemistry). The rate of returning to the first state, RISC in target finding mode is represented by k_3. Note: that the reverse reaction of cleavage, ligation, is unreported/unknown with Argonaute proteins, this is denoted by k_{-2} and the thin arrow.

Table S3.1: RNAs and DNAs used in this study.

Synthetic guide strand description	**Sequence** Passenger, guide, DNA guide, seed, m indicates 2′-O-methyl ribose; p indicates 5′ monophosphate Bio, Biotin-6-carbon spacer; *, phosphorothioate
Passenger strand for unlabeled let-7a RISC	pUAU ACA ACC UAC UAC CUC CUU
Guide strand for unlabeled let-7a RISC	pUGA GGU AGU AGG UUG UAU AGU
Passenger strand for Alexa let-7a RISC	mUAU ACA ACC UAC UAC CUG CUU
Guide strand for Alexa let-7a RISC	pUGA GGU AGU AGG UUG UAU AGU-NH_2
DNA guide strand with first 16 nt of let-7a	pTGA GGT AGT AGG TTG T
DNA guide strand with first 16 nt of let-7a for Alexa labeling	pTGA GGT AGT AGG TTG T-NH_2
Passenger strand for Alexa miR-21 RISC	mACA UCA GUC UGA UAA GCA UUU
Guide strand for Alexa miR-21 RISC	pUAG CUU AUC AGA CUG AUG UUG-NH_2
28 nt let-7a perfect RNA target	Bio-GAU ACU AUA CAA CmC*mU ACU ACC UCA ACC U-NH_2
28 nt let-7a seed only target	Bio-GAA AAA AAA AAA AAA UCU ACC UCU AAA U-NH_2
Synthetic guide strand description	**Sequence** RNA, DNA; m, 2′-O-methyl; p, 5′ phosphate; Bio, Biotin-6-carbon spacer; RISC-binding site; 2′-N-U, Uridine with 2′ amine; Dap, 2,6 diaminopurine
Capture Oligo to affinity purify let-7a RISC	Bio-mAmUmA mGmAmC mUmGmC mGmAmC mAmAmU mAmGmC mCmUmA mCmCmU mCmCmG mAmAmC mG
DNA competitor to elute let-7a RISC	CGT TCG GAG GTA GGC TAT TGT CGC AGT CTA T-Bio
Capture Oligo to affinity purify miR-21 RISC	Bio-mGmAmU mGmAmA mCmCmA mCmUmC mAmGmA mGmAmC mAmUmA mAmGmC mUmAmA mUmCmU mA
DNA competitor to elute miR-21 RISC	TAG ATT AGC TTA TGT CTC TGA GTG GTT CAT C-Bio

Forward primer to generate templates for T7 transcription of let-7a target RNAs	GCG TAA TAC GAC TCA CTA TAG GGT TTT AAT GAA TAC GAT TT
Reverse primer to generate template for T7 transcription of fully complementary let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of seed-matched let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG TAG ATC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of target with seed match plus 3′ supplementary pairing to let-7a	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG TAG ATC CTT GTT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of t2t3 seed-mismatched let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTC TGG TAG ATC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of t3t4 seed-mismatched let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG TCG TAG ATC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of t4t5 seed-mismatched let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG ATT TAG ATC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of t5t6 seed-mismatched let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGC AAG ATC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of t6t7 seed-mismatched let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG ATG ATC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of t7t8 seed-mismatched let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG TTC ATC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G

Reverse primer to generate template for T7 transcription of t8 seed-mismatched let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG TAC ATC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of seed + t9 let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG TAG TTC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of seed + t9t10 let-7a target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG AGG TAG TAC CAA CAT ATC AAT CCA GAG GAA TTC ATT ATC AGT G
let-7a 3′ product mimic with seed matched paring to target	pCUA CCU CAA CCU UUU AUA CAC AGU UU(2'-N-U) CCG-Bio
let-7a 3′ product mimic with seed matched +t9 pairing to target	pACU ACC UCA ACC UUU UAU ACA CAG UU(2'-N-U) CCG-Bio
let-7a 3′ product mimic with seed matched +t9t10 paring to target	pUAC UAC CUC AAC CUU UUA UAC ACA GU(2'-N-U) CCG-Bio
Reverse primer to generate template for T7 transcription of fully complementary let7a target except for a t4t5 seed-mismatch	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTG ATT TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of non-complementary target	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTT AGC TAT AAT GAA ATG CCT TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer to generate template for T7 transcription of target complementary to 3′ half of let-7a	CCC ATT TAG GTG ACA CTA TAG ATT TAT ACC TAG TTA AAC AGC GGA ACT GTG TAT AAA AGG TTT TTT TTT TTG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
5′-biotin-labeled, fully complementary let-7a DNA target	Bio-GGG TTT TAA TGA ATA CGA TTT TGT ACC AGA GTC CTT TGA TCG TGA CAA AAC AAT TGC ACT GAT AAT GAA TTG GTC TGG ATA CTA TAC AAC CTA CTA CCT CAA CCT TTT ATA CAC AGT TCC GCT GTT TAA CTA GAT GT

5′-biotin-labeled, seed-matched let-7a DNA target	Bio-GGG TTT TAA TGA ATA CGA TTT TGT ACC AG AGT CCT TTG ATC GTG ACA AAA CAA TTG CAC TGA TAA TGA ATT CCT CTG GAT TGA TAT GTT GGA TCT ACC TCA ACC TTT TAT ACA CAG TTC CGC TGT TTA ACT AGA TGT
DNA target with 5′-biotin and seed-match plus 3′ supplementary pairing to let-7a	Bio-GGG TTT TAA TGA ATA CGA TTT TGT ACC AGA GTC CTT TGA TCG TGA CAA AAC AAT TGC ACT GAT AAT GAA TTC CTC TGG ATT GAT AAC AAG GAT CTA CCT CAA CCT TTT ATA CAC AGT TCC GCT GTT TAA CTA GAT GT
Synthetic DNA-RNA for attaching Alexa DNA piece to 5′ end of IVT targets	pTGT CAC CTA GAT CGA UGA AUU CCU CUG GAU CAC ACA CAA AAA AAA
Splint for DNA-RNA ligation	TAT TCA TTA AAA CCC TTT TTT TTG TGT GTG
Splint for ligation of Alexa647 DNA to 5′ end of RNA target (3′ bound target)	CTA GGT GAC AGT GGT TGG GT
Forward primer for miR-21 perfect target	GCG TAA TAC GAC TCA CTA TAG GGG TCC TTT GAT CGT GAC AAA ACA AT
Reverse primer for miR-21 perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC TAG TTG AGG TGC GGA ACT GTG TAT AAA AGG TTA GCT TAT CAG ACT GAT GTT GAA TCC AGA GGA ATT CAT TAT CAG TG
miR-21 seed only target	Bio-GAU AAA AAA AAA AAA AAU AAG CUA ACC U.NH_2
miR-21 seed only containing 2,6 Diaminopurine target	Bio-GAU AAA AAA AAA AAA ADapU DapDapG CUA ACC U.NH_2
miR-21 seed + 3′ supplementary target	Bio-GAU AAA AAU CAG AAA AAU AAG CUA ACC U.NH_2
RNA oligo for La binding	P.CGC GGC GCC UAC UAC CUC UUU
Gapmer to direct 3′ end trimming with RNase H	mCmAmC mUmAmU mAmGA TTT mAmCmA mUmCmU mAmG

Substrates	**Sequence** RNA, DNA; BioG, 5′-Biotin-G; Bio, Biotin-6-carbon spacer; U, Alexa Fluor 647 deoxyuridine; target site/pairing to RISC 2′-N-U, Uridine with 2′ amine; Dap, 2,6 diaminopurine p, 5′ phosphate
Klenow polymerase template to synthesize 3′ DNA extension containing 17 Alexa Fluor 647 dyes	ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT TAT TTA CAT CTA GTT GAG GTG CGG AAC TG
Trap oligonucleotide for the preceding template (fully complementary)	CAG TTC CGC ACC TCA ACT AGA TGT AAA TAA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA AT
Klenow polymerase template for the synthesis of a short 3′ DNA extension with a single biotin	GTT TAT TTA CAT CTA GTT GAG GTG CGG AAC TG
Klenow polymerase template for the synthesis of single-stranded DNA extension containing 17 Alexa Fluor 647 dyes	ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT ATT GTT GTT AGG GTT TTA ATG AAT ACG ATT TTG TAC CAG AGT CC
DNA primer for the preceding template	GGA CTC TGG TAC AAA ATC GTA TTC ATT AAA ACC C
Trap oligonucleotide for the preceding template (fully complementary)	GGA CTC TGG TAC AAA ATC GTA TTC ATT AAA ACC CTA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA ATA ACA ACA AT
5′-tethered, target with complete complementarity to a let-7a and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	Bio-GGG UUU UAA UGA AUA CGA UUU UGU ACC AGA GUC CUU UGA UCG UGA CAA AAC AAU UGC ACU GAU AAU GAA UUC CUC UGG AUA CUA UAC AAC CUA CUA CCU CAA CCU UUU AUA CAC AGU UCC GCU GUU UAA CUA GAU GUA AAU AAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA U

5′-tethered, target with a let-7a seed-match and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AUC UAC CUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered, target with a let-7a seed-match plus 3′ supplementary pairing and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA ACA AGG AUC UAC CUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG AUG UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered, non-complementary target with 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UAA GGC AUU UCA UUA UAG CUA AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered target with a let-7a seed-match but for a t2t3:g2g3 mismatch and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AUC UAC CGA AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered target with a let-7a seed-match but for a t3t4:g3g4 mismatch and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AUC UAC AGC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU

5′-tethered target with a let-7a seed-match but for a t4t4:g4g5 mismatch and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AUC UAA AUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered target with a let-7a seed-match but for a t5t6:g5g6 mismatch and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AUC UUG CUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered target with a let-7a seed-match but for a t6t7:g6g7 mismatch and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AUC AUC CUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered, target containing a seed-match to let-7a but for a t7t8:g7g8 mismatch and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AUG UAC CUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered, target containing a seed-match to let-7a but for a t8:g8 mismatch and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AUG UAC CUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU

5′-tethered, target with a let-7a seed-match +t9 and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG AAC UAC CUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered, target with a let-7a seed-match +t9t10 and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUG AUA UGU UGG UAC UAC CUC AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG GUA UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
let-7a 3′ product mimic with seed matched paring to target	pCUA CCU CAA CCU UUU AUA CAC AGU UU(2'-N-Alexa647-U) CCG-Bio
let-7a 3′ product mimic with seed matched +t9 pairing to target	pACU ACC UCA ACC UUU UAU ACA CAG UU(2'-N- Alexa647-U) CCG-Bio
let-7a 3′ product mimic with seed matched +t9t10 paring to target	pUAC UAC CUC AAC CUU UUA UAC ACA GU(2'-N-Alexa647-U) CCG-Bio
5′-tethered, target with a complete complementarity to let-7a except for a dinucleotide mismatch to the let-7a seed and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	Bio-GGG UUU UAA UGA AUA CGA UUU UGU ACC AGA GUC CUU UGA UCG UGA CAA AAC AAU UGC ACU GAU AAU GAA UUC CUC UGG AUA CUA UAC AAC CUA CUA AAU CAA CCU UUU AUA CAC AGU UCC GCU GUU UAA CUA GGU AUA AAU AAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA U

5′-tethered, target with complementarity to the 3′ half of let-7a and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGU UUU AAU GAA UAC GAU UUU GUA CCA GAG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UAC UAU ACA ACC AAA AAA AAA AAC CUU UUA UAC ACA GUU CCG CUG UUU AAC UAG AUG UAA AUA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered, DNA target with complete complementarity to let-7a and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	Bio-GGG TTT TAA TGA ATA CGA TTT TGT ACC AGA GTC CTT TGA TCG TGA CAA AAC AAT TGC ACT GAT AAT GAA TTG GTC TGG ATA CTA TAC AAC CTA CTA CCT CAA CCT TTT ATA CAC AGT TCC GCT GTT TAA CTA GAT GTA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered, DNA target with let-7a seed-match and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	Bio-GGG TTT TAA TGA ATA CGA TTT TGT ACC AGA GTC CTT TGA TCG TGA CAA AAC AAT TGC ACT GAT AAT GAA TTC CTC TGG ATT GAT ATG TTG GAT CTA CCT CAA CCT TTT ATA CAC AGT TCC GCT GTT TAA CTA GAT GTA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
5′-tethered, DNA target with let-7a seed-match plus 3′ supplementary pairing and a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	Bio-GGG TTT TAA TGA ATA CGA TTT TGT ACC AGA GTC CTT TGA TCG TGA CAA AAC AAT TGC ACT GAT AAT GAA TTC CTC TGG ATT GAT AAC AAG GAT CTA CCT CAA CCT TTT ATA CAC AGT TCC GCT GTT TAA CTA GAT GTA AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU AAC AAC AAU
Single-stranded DNA containing 17 Alexa Fluor 647 dyes used to make 3′-tethered targets	CGT AGA CGC TCT TTT CAG CCA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UCC CAC CCC CAA ACC CAA CCA C

3′-tethered, fully complementary let-7a target containing a 5′ single-stranded DNA extension containing 17 Alexa Fluor 647 dyes (added by ligation)	CGT AGA CGC TCT TTT CAG CCA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UCC CAC CCC CAA ACC CAA CCA CTG TCA CCT AGA TCG AUG AAU UCC UCU GGA UCA CAC ACA AAA AAA AGG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU ACU AUA CAA CCU ACU ACC UCA ACC UUU UAU ACA CAG UUC CGC ACC UCA ACU AGA UGU AAA TAA AC-Bio
5′-tethered, fully complementary miR-21 target containing a 3′ DNA extension containing 17 Alexa Fluor 647 dyes	BioG-GGG UCC UUU GAU CGU GAC AAA ACA AUU GCA CUG AUA AUG AAU UCC UCU GGA UUC AAC AUC AGU CUG AUA AGC UAA CCU UUU AUA CAC AGU UCC GCA CCU CAA CUA GAU GUA AAU AAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA U
miR-21 seed only target	Bio-GAU AAA AAA AAA AAA AAU AAG CUA ACC U.NH-Alexa647
miR-21 seed only containing 2,6 Diaminopurine target	Bio-GAU AAA AAA AAA AAA ADapU DapDapG CUA ACC U.NH-Alexa647
miR-21 seed + 3′ supplementary target	Bio-GAU AAA AAU CAG AAA AAU AAG CUA ACC U.NH-Alexa647
3′-tethered, fully complementary miR-21 target containing 5′-ligated single-stranded DNA extension containing 17 Alexa Fluor 647 dyes	CGT AGA CGC TCT TTT CAG CCA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UAA CAA CAA UCC CAC CCC CAA ACC CAA CCA CTG TCA CCT AGA TCG AUG AAU UCC UCU GGA UCA CAC ACA AAA AAA AGG GGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU UCA ACA UCA GUC UGA UAA GCU AAC CUU UUA UAC ACA GUU CCG CAC CUC AAC UAG AUG UAA AUA AAT AAA C-Bio

CHAPTER IV: Perspectives and Future Work

Summary

The Argonaute protein is conserved in all kingdoms of life and it is the core effector protein in small regulatory RNA pathways. The study of Argonaute enzymology gives us insight into its regulatory roles inside cells but also may tell us something about how RNA-guided endonucleases have evolved. The Argonaute protein has distinct kinetic features, which it obtains from changing the biophysical properties of its bound small RNA guide. These properties correlate with the functional role of Argonaute in a given organism or tissue types. We sought to understand the initial and terminal steps in small RNA regulation in mammals and we have learned how Argonaute proteins have specific kinetic features related to their function. This thesis highlights the insights we gained with Argonaute and how it can change the biophysical properties of nucleic acid. This work is a stimulus to further understand small RNA pathways and may reflect a general mechanism of RNA-guided endonucleases to which future studies can draw parallels against.

It's a duck! No, it's a plane! No, it is AGO-RISC finding its target!

In Chapter III, we directly measured the association rate (k_{on}) of mouse AGO2-RISC and *Thermus thermophilus* AGO bound to a small DNA (Figure 3.2 and Table 3.1). The ability for Argonaute proteins to change the rate of nucleic acid hybridization had been previously inferred from structural studies and

derived from in vitro ensemble kinetic studies (Ma et al., 2004; Parker et al., 2005; Wang et al., 2008a; Parker et al., 2009; Elkayam et al., 2012; Nakanishi et al., 2012; Wee et al., 2012). Changing the rate of nucleic acid hybridization is a fundamental property of the small regulatory RNA pathway since nucleic acid hybridization is a slow process and not efficient under physiologic conditions (Herschlag, 1991; Herzog and Ameres, 2015). A comparison of the association rates (k_{on}) of the small RNA guide alone versus the small RNA guide bound to an Argonaute protein shows an acceleration of as much as 250-fold (Figures 3.1 and 3.2). Both nucleic acid hybridization and RISC target finding use three-dimensional diffusion in order to find and bind to their target. Thus, both reactions can be influenced by temperature and viscosity; however, nucleic acid hybridization is not a diffusion-controlled reaction because of its freely rotating nucleotides. As explained in Chapter I and illustrated in Figure 1.6, nucleic acids have an entropic penalty associated with hybridization. Rotation of the individual nucleotides combined with potential intra- and inter-molecular interactions lead to many non-productive collisions and therefore an overall on-rate that is slower than diffusion limits. Even though a 21 nt RNA (~7 kDa) is only a fraction of the molecular weight of RISC — the 21 nt RNA plus a ~100 kDa protein — the k_{on} for RISC is faster (Figure 3.1). This acceleration is attributed to Argonaute pre-paying the entropic penalties associated with nucleic acid hybridization by pre-organizing a sub-set of the bound small RNA guide to form a pre-helical

shape (Ma et al., 2004; Parker et al., 2005; Wang et al., 2008a; Parker et al., 2009; Elkayam et al., 2012; Nakanishi et al., 2012; Schirle and MacRae, 2012; Faehnle et al., 2013). This structural observation has been supported by kinetic studies that inferred k_{on} and these values are now confirmed by our direct measurements (Figure 3.2 and Table 3.1; Wee et al., 2012; Deerberg et al., 2013). The protein cradles and anchors this portion of the small RNA guide (g1 and the seed, g2–g8), Figures 1.4 and 1.5) so that it no longer acts like a nucleic acid but more like a RNA-binding domain of a protein. The 'genius' of Argonaute is that seed is like a modular binding domain that is programmed by the bound small RNA. This feature allows Argonaute proteins to have the ability to regulate many RNAs and find those targets with nucleotide precision.

Another perspective to the small regulatory RNAs that relates to nucleic hybridization and entropic penalties is the biogenesis of RISC that typically starts from a double-stranded pre-cursor and NOT a single-stranded one. This is important for nuclease stability inside the cell, but more importantly, double-stranded RNA overcomes potential intra- and inter-molecular interactions that single-stranded RNA would make in the cell thus preventing RISC formation. The siRNA or miRNA duplex is key to AGO being able to select the guide strand or miRNA strand and efficiently assemble it into the protein. The revised RISC assembly model proposed by the Tomari lab lends credence to this thought since the chaperones keep AGO 'open for business' and allow

binding of the RNA duplex (Kobayashi and Tomari, 2015). After forming pre-RISC, the chaperones are no longer required and Argonaute is able to remove one strand to form the RISC with no additional co-factors (Kobayashi and Tomari, 2015). This also explains why recombinant Argonaute cannot form RISC with a duplex siRNA but rather needs a single-stranded RNA to form RISC (Rivas et al., 2005; MacRae et al., 2008). The protein without a small RNA bound is either unstable or potentially in a more energetic state that is not conducive for binding a RNA duplex (Johnston et al., 2010; Kobayashi and Tomari, 2015). Is it possible that the current structural information that is generated by forming RISC with a single-stranded guide RNA and recombinant protein misinterpret how the guide is positioned in vivo? Future structural studies of affinity purified RISC that was assembled in cellular extracts or loaded via chaperone protein reconstitution along with recombinant AGO may help us understand if any differences exist. Intriguingly, this brings up quandary for small RNA regulation since the PIWI clade Argonautes forms RISC via single-stranded RNA. How is this possible and/or efficient? How does this pathway overcome the inherent problem with assembling single-stranded RNAs? Do RNA-binding proteins exist that are associated with the biogenesis machinery and take the place of not having a complementary RNA that assists with proper PIWI-RISC formation? Future studies to create a biochemical reconstitution extract for piRNAs and piRNA biogenesis can hopefully answer this question along with many others.

Dwelling on dwell times

It is remarkable that RISC bound to a seed only target can decrease the dissociation rate (k_{off}) by ~14,000-fold when compared to the same target paired via seven base pairs of a RNA guide strand alone (Figure 3.3 and Table 3.1). Simply put, AGO2-RISC remains bound to a RNA target that only contains seven nucleotides of base pairing for nearly 5 minutes at 37°C versus the fractions of a second for RNA alone (Figure 3.3). Our results agree well with our previous ensemble filter-binding measurements at room temperature (Wee et al., 2012). Future studies combining TNRC6A (GW182) along with other RNA regulatory factors may reveal that this dwell time increases so that destabilization and/or translational repression of the RNA target can take place. Furthermore, future work in this area will help to address fundamental questions related to how miRNA regulate their targets and maybe settle the long-standing debate in the field regarding mRNA destabilization and/or translation repression as the mode of action.

While mammalian AGO2 remains bound to seed paired targets for minutes, *Thermus thermophilus* AGO and fly AGO2 only remain bound for only seconds (Table 3.1; Wee et al., 2012). The k_{on} for both these Argonaute proteins is similar to mammalian AGO2 but why is the mean dwell time so short? Could this reflect the evolved functions that each of these Argonautes play inside the cell? Argonaute proteins regulate in two manners, either by slicing or by binding

to create a scaffold for regulatory factors (Figure 1.1). Thus, we can segregate the Argonautes by their function: slicer or non-slicer. Furthermore, there are Argonautes that only slice to regulate (*e.g.,* fly AGO2), let's call them 'professional' slicers and then there slicer Argonautes that can do both binding and slicing methods of regulation (*e.g.,* mammalian AGO2), let's call them 'dual-mode' slicers. In mammals, there is no strict division/sorting of small RNAs among the four AGO proteins, since catalytically active or not, each participates in the miRNA regulatory pathway (Hutvagner and Simard, 2008). Mammalian miRNAs typically do not pair perfectly to their targets; therefore, they never need nor can they slice, rather they act as binding scaffolds (Jonas and Izaurralde, 2015). The one exception is mammalian AGO2 that is a 'dual-mode' slicer. Primarily, AGO2 binds miRNAs and regulates through a non-slicing pathway but it proposed that the retained catalytic activity is required for processing specific dicer-independent miRNAs (*e.g.,* mir-451) and plays a role in embryonic development (Cheloufi et al., 2010). Meanwhile, 'professional' slicers have a cellular function to regulate targets with perfect complementarity to their guides. This is clear in flies that have small RNA sorting (Ghildiyal et al., 2010); a clear example is fly Ago2 which has a role in viral defense (Wang et al., 2006) and it is kinetically advantageous for Ago2-RISC not to get 'distracted' by targets with partial complementarity or seed only pairings that would prevent slicing. On the other hand, fly Ago1, the miRNA-specific Argonaute, binds targets with partial

complementarity (seed and/or seed + 3′ supplementary pairing) and regulates via mRNA destabilization/translation repression. The affinity (K_D) to targets for 'professional' slicers versus non-slicers or 'dual-mode' slicers have distinctive differences. Mammalian AGO2 (a dual-mode slicer) and fly Ago1 (Flores-Jasso and Zamore, unpublished) have similar affinities for seed only, seed + 3′ supplementary, and perfectly paired targets whereas fly AGO2, a professional slicer, has a clear preference for perfect targets (Wee et al., 2012). Fly AGO2′s affinity for perfect targets is ~4 pM whereas as its K_D to seed only paired targets is ~200 pM (Wee et al., 2012). Another example of the binding affinity difference between 'professional' and 'non-professional' slicers is seen with TtAgo. There are no known miRNAs in *Thermus thermophilus* but there are small DNAs that are hypothesized to have a role in the defense against foreign plasmids (Swarts et al., 2014a). The affinity of TtAgo bound with a small DNA to a perfect target is ~ 10 pM (Jolly and Zamore, unpublished) but similar to fly AGO2, there is a large difference to partially paired targets where seed only targets have a K_D ~20 nM (Table 3.1). The common theme among all of these Argonautes is that the affinity to a specific target is congruent with its function. The 'professional' slicing Argonautes have a greater affinity toward perfect or highly complementary targets and Argonautes that primarily regulate without slicing have equivalent affinities to any target that contains a miRNA-like pairing or perfect pairing.

Seeking out a seed

Beyond target finding and binding is the remarkable selectivity of Argonautes have at distinguishing partial seed pairing versus full seed paring. We performed a series of experiments with seed only targets that contained dinucleotide mismatches starting from t2 going through end of the seed, t8 (Figure 3.3). Nucleic acid hybridization rules would dictate that central mismatches are expected to be the most detrimental as they would break the half turn of the helix and terminal mismatches would have a less deleterious effect (Figure S3.7). Instead, we observed that the seed itself contains sub-domains, the 5′ end sub-domain of the seed with more importance for the fast k_{on} and it confers the majority of the stabilizing interaction with the target, whereas the 3′ end sub-domain contributes little to no acceleration for k_{on} but is important for additional stability and potentially completes protein conformation changes that create the stable RISC-target complex (Figures 3.3, S3.7, and 4.1).

The 5′ sub-domain of the seed, nucleotides g2–g5, acts like a cooperative binding unit, where any mismatch within this sub-domain has the same impact on affinity (Figures S3.7 and 4.1). A full seed pairing has a K_D ~15 pM whereas seed mismatches in the first domain have a K_D ~ 250 nM. The ~17,000-fold lower affinity for a miRNA target with mismatches in the 5′ sub-domain illustrates one of the properties of Argonaute to discriminate real targets from close matches. The transition point between the 5′ sub-domain

correlates nicely with a recent structural study from Schirle and colleagues, the structure of human AGO2-RISC bound to a short miRNA target (Schirle et al., 2014). This structure explained a previous observation that α-helix 7, found in the L2 linker between the PAZ and MID domains, would clash and create a kink between AGO2-RISC pairing to a target (Figure 1.4, top panel; Schirle and MacRae, 2012; Schirle et al., 2014). Comparison of the two structures, target bound and unbound, showed that α-helix 7 moves ~4 Å and by doing so removes the kink or interruption in the helix (Figure 1.4; Schirle et al., 2014). The movement of α-helix 7 and change in protein conformation may represent the mechanism by which RISC interrogates and discriminates targets (Schirle et al., 2014; Swarts et al., 2014b). Complete seed base pairing must start from the 5′ end to help with proper change in protein conformation to accommodate additional base pairs g6–g8. Our kinetic measurements agree well with the structural data that a transition point exists in the seed and therefore creates sub-domains. Furthermore, our single-molecule work presented in Chapter III along work published at the same time from the Joo and MacRae labs come to similar conclusions; a subset of the seed is used to interrogate targets (Chandradoss et al., 2015; Salomon et al., 2015).

Figure 4.1

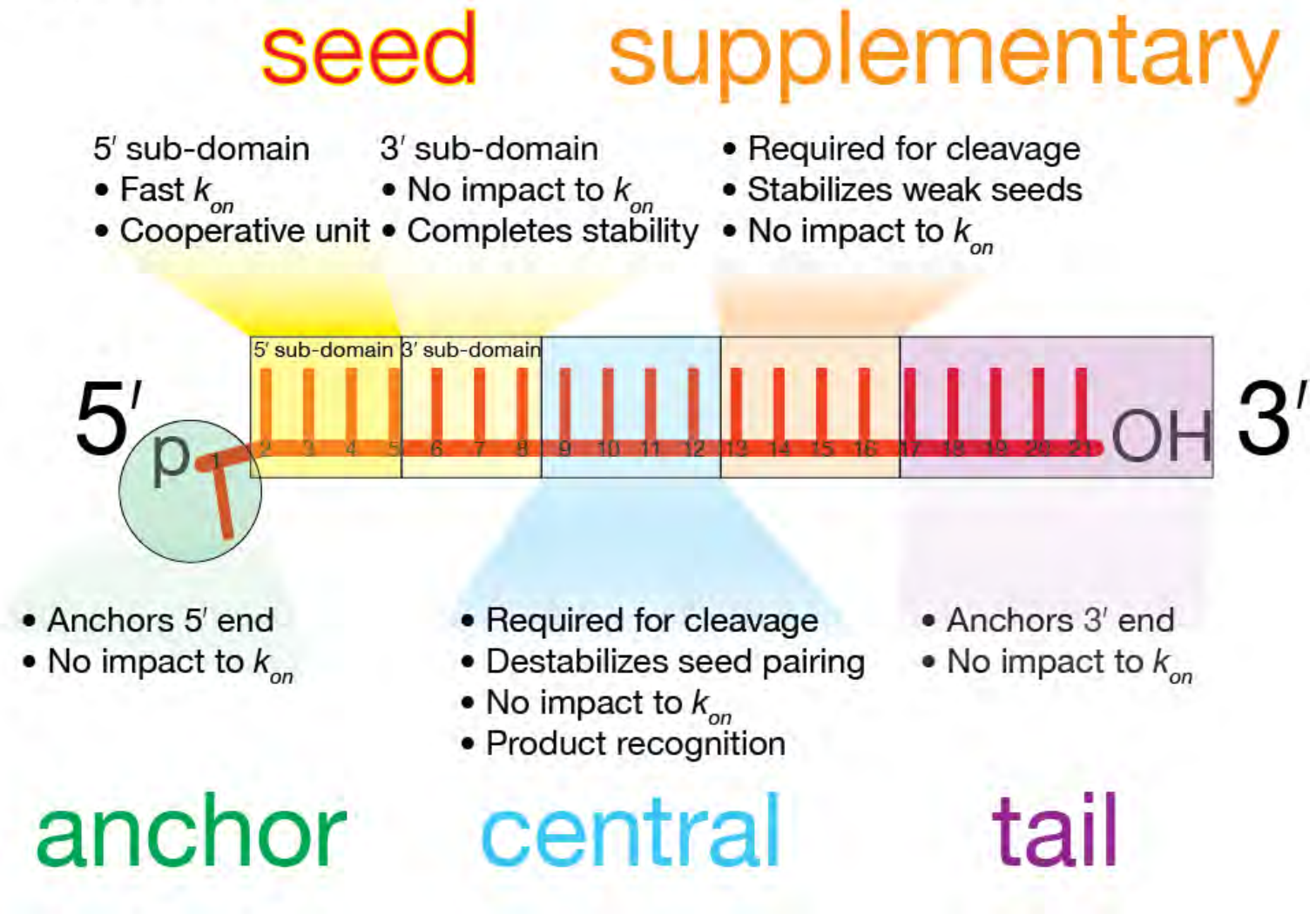

seed
supplementary
5′ sub-domain
• Fast k_{on}
• Cooperative unit
3′ sub-domain
• No impact to k_{on}
• Completes stability
• Required for cleavage
• Stabilizes weak seeds
• No impact to k_{on}
5′ sub-domain
3′ sub-domain
5′
p
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
OH
3′
• Anchors 5′ end
• No impact to k_{on}
• Required for cleavage
• Destabilizes seed pairing
• No impact to k_{on}
• Product recognition
• Anchors 3′ end
• No impact to k_{on}
anchor
central
tail

Figure 4.1: Revised functional domain map of the small RNA guide.

Argonaute changes the shape of the small RNA guide when it is in a complex with it. Distinct functional domains are created that will change the biophysical properties of the bound small RNA. This map reflects the newly discovered sub-domains of the seed while pointing out important features of that were previously described (Wee et al., 2012).

The 3′ sub-domain of the seed provides more than just additional base pair stability between RISC and the target (Figure 4.1). Target mismatches to the 3′ sub-domain of the seed do correspond to the *ΔΔG* penalties that are predicted by the Nearest-neighbor model BUT, the overall seed pairing stability is not just coming from the 5′ end sub-domain. There is interplay of Watson-Crick base pairs in a complete seed pairing along with interactions that can't just be assigned to base pairing to the target alone. Another conformation change when AGO2 binds a complete seed (g2–g8 paired to t2–t8) is the PAZ domain turning (Schirle et al., 2014). The PAZ domain along with α-helix 7 moves like a rigid body or 'hinge' along the other domains and is seen to open or widen the N-PAZ channel that may facilitate protein-target interactions (Schirle et al., 2014). Future structural studies with longer targets that have seed only pairing will help us understand how Argonaute increases its affinity to a miRNA site ~17,000 times tighter than nucleic acid pairing alone.

Argonautes prefer substrates over products

When RISC is bound to a miRNA target through a full seed pairing it forms a fairly stable complex that presumably allows it to regulate the target by acting as a scaffold for the destabilization/translation repression machinery. What would happen if Argonaute were bound to a target that had pairing past the seed? One could *reasonably* predict it should be more stable. There are a sub-set of miRNA targets that have pairing beyond the seed called seed + 3′ supplementary

paired targets (pairing t2–t8, t13–t16; Bartel, 2009); Why did these miRNAs not evolve to have contiguous base pairs through the central region? Especially since many miRNA Argonautes do not contain residues or proper domain structure that would allow them to be catalytically active? For example, miR-21 has a thermodynamically weaker seed than let-7a, ΔG= –13.3 versus –15.6, respectively. When miR-21-RISC pairs to a seed + 3′ supplementary target it is more stable than when it is on a seed only target whereas for let-7a there is little to no difference between these types of pairing (Figures 3.3B and S3.5D). Although, we did not originally set out to answer this question, an important insight was discovered when we measured the kinetics of RNAi product release. AGO2-RISC has a high affinity to miRNA targets which bear seven nucleotides of complementarity, however, counterintuitive to nucleic acid base pairing rules, the affinity decreases when RISC is paired one or two additional nucleotides past the seed (seed + t9 (8 nt) or seed + t9t10 (9 nt) pairing, Figure 3.8). The affinity (K_D) goes from 15 pM for a seed only target to 90 pM for the seed plus one additional base pair and it is 50 pM for the seed plus two additional base pairs. We initiated direct measurement of this type of pairing based on our results measuring product release after RNAi where AGO2-RISC left the 3′ cleavage product (seed + t9 and t10) ~14 times faster than if it were bound to a seed only target. The additional two nucleotides with the cleavage product were actually *destabilizing* instead stabilizing. To further understand the difference

and potential influence of having a full-length RNA over a cleavage product or ‘stump’ led to a panel of different targets (Figure 3.8). Are there RNA-protein interactions that help Argonaute recognize a substrate versus product? We tested 3 different ‘stumps’ that mimicked cleavage products and varied in base pairing (7–9 nt) but the total length (30 nt) did not change. Again, AGO2-RISC does not follow the rules of nucleic acid pairing, additional base pairs that *should* increase the mean dwell time (tau, τ), were about the same for 7–9 nt of base pairing (Figure 3.8, τ ~30–50 seconds). Re-examination of our previous results with the mismatches in the 3′ end seed sub-domain dissociate faster than the stumps alone, and those targets were in the context of a full-length target. This argues that it is not just about having RNA extending past the point of pairing rather it is the amount of pairing that may confer a protein conformation that dictates the interactions with the target and it cannot just be attributed to Watson-Crick base pairing. Stability is not only about the number of base pairs rather the order and position of base pairs. For example, AGO2-RISC that has no pairing to the 5′ end of the guide (*i.e.*, no seed pairing, t11–t21 target) has a slow k_{on} (Figure 3.2B) and is close to what is measured for nucleic acid alone, $3.6 \pm 0.2 \times 10^{7}$ $M^{-1} \cdot s^{-1}$ for AGO2-RISC and $7.6 \pm 0.9 \times 10^{6}$ $M^{-1} \cdot s^{-1}$ for the naked guide strand. Furthermore, the mean dwell time for AGO2-RISC paired to the t11–t21 target is also similar to nucleic acid alone, τ ~1.3 seconds versus τ ~1.7 seconds (predicted) (Figures 3.2B and 3.3B). This idea that RISC

stability is a combination of base pair position and target interaction will need further kinetic studies to figure out how target length plays a role. Measurements of the mean dwell time of AGO2-RISC on different lengths of target that keep a constant amount of Watson-Crick base pairs (*i.e.,* seed only pairing with different target lengths) could address this question. The experimental setup would mimic our measurements with the cleavage product mimics in Figure 3.8 but systematically extend the RNA at the 5′ end toward the 3′ end of the guide strand. Combinations of kinetic studies along with a structural analysis may illustrate an elegant interplay of base pairing and non-base pair interactions of RISC and the target.

Seed strength and slicing

Our analysis of product release and global fitting of the rate of target finding and cleavage/product release measured a lag between target binding and product release (Figure 3.6) that shows slow step in the RNAi reaction step. What molecular events lead to this lag between substrate binding and product release (Figure 3.6, *k* represents measured rate)? We cannot measure when the target is cleaved by nucleophilic attack of the scissile phosphate across from g10 and g11 of the guide but we can postulate why there is a lag between binding a product release (Figure S3.8A, *k*), assuming the chemistry step is not rate limiting (Figure S3.8C, k_2). Previous structural studies have revealed that when

RISC is unbound to a target it is not in a cleavage competent state since the catalytic residues that coordinate the Mg^{2+} ion are not positioned correctly (Wang et al., 2008a; Wang et al., 2008b; Wang et al., 2009; Sheng et al., 2014). This means there needs to be a two-step binding process where the target is bound then is followed by a conformation change that allows slicing to generate the $E{\cdot}P_1{\cdot}P_2$ complex (Figures 3.6B, S3.8A, and S3.8C). The model for having two conformation states was originally proposed by Tomari and Zamore; the "two-state" model (Tomari and Zamore, 2005). What can influence the rate of changing from one state to the other? Does the transition go back and forth as the protein negotiates the target to correctly position it for cleavage? We measured that stronger seeds like let-7a have a slower *k* when compared to weaker seeds like miR-21, $0.15 \pm 0.01\ s^{-1}$ versus $0.31 \pm 0.01\ s^{-1}$, respectively. Why is this rate about two times slower for let-7a when compared to miR-21? Can this be just attributed to seed strength? Our measurements with a target only containing seed mismatches at t4t5 shows a slight increase in the rate but does not directly address these questions; furthermore the mismatches may have their own impact on conformation changes and would not be the best approach to answer this question. Recently, a single-molecule FRET study with the archaeal Argonaute from *Methanocaldococcus jannaschii* (MjAgo) shows two FRET states and concluded that this would support the 'two-state' model (Zander et al., 2014). The biology of MjAgo is not well understood but there are

no known miRNAs in *M. jannaschii* and it is a catalytically active Argonaute (Makarova et al., 2011; Zander et al., 2014). MjAgo is likely to be a 'professional' slicer Argonaute where it has a higher affinity for perfect targets over partially complementary targets (*e.g.,* seed only pairing), similar to fly AGO2 and TtAGO. If the same FRET measurements are made with mammalian AGO2 will there be a similar transition? The slow transition state for AGO2 may resemble a safety switch since its primary cellular role is NOT to cleave targets but rather regulate through the binding (Figure S3.8C, k_1). The requirement for a rigid transition state would ensure fidelity in slicing, especially when it is not the primary mode of regulation. Furthermore, can the strength of the seed relate back to the potency of siRNAs in vitro and in vivo? Massive screens performed by those developing therapeutic siRNAs and RNAi reagents (Reynolds et al., 2004) may be able to explain the added potency observed with some sequences versus others by relating back the seed strength pairing to a dwell time that reflects conformation changes, cleavage, and product release. Future studies dissecting known 'hyper-functional' siRNAs versus poor siRNA sequences that are independent of RNA target secondary structure and/or RISC loading may aid in creating rules for the better design of siRNAs.

Argonaute: Nature's Programmable Regulator

Argonaute participates in several regulatory functions in all kingdoms of life (Ameres and Zamore, 2013; Swarts et al., 2014b). A primary function in

mammals is related to post-transcriptional gene regulation but there is evidence for a role in the nucleus for DNA-damage response and gene activation (Meister, 2013). The work presented in this thesis measures and further develops our understanding to the molecular functions of the Argonaute protein bound to a small RNA or DNA guide. The nucleic acid guide no longer has properties like protein-free nucleic acids. The association rate (k_{on}) is accelerated and the dissociation rate is changed through a potential interplay of protein-nucleic acid interactions that are determined by the combination of base pairing position of the guide to the target. RNA-guided endonucleases in general can be looked at as regulatory factors that can quickly adapt to the changing environment of the cell either by quickly counteracting a viral invader sequence or specifically regulating a transcript during a temporal cellular event. Unlike transcription factors or RNA-binding proteins, Argonaute did not need to evolve a specific protein-based RNA-binding domain; instead Ago converts a bound RNA into specific functional domains (Wee et al., 2012), one being the seed, which is the equivalent of a binding domain. Nucleic acid pairing acceleration is a common attribute of all Argonautes and our measurements show it is conserved over 2 billion years of evolutions between mice and the thermophilic eubacterium, *Thermus thermophilus*. There appears to be two major functional classes of Argonaute proteins, slicers and non-slicers. This distinction is not just about having the catalytic residues to cleave a nucleic acid targets but more a

combination of binding properties coupled to catalytic ability. There are ‘professional’ slicers like TtAgo or fly AGO2 that cleave in vivo AND have binding affinities for perfect targets over partially complementary targets. There are ‘dual-mode’ Argonautes like mouse and human AGO2 that can do both forms of regulation slicing and miRNA-based regulation. The ‘dual-mode’ of regulation is not fully understood and future structurally studies may reveal that there are specific interactions that take place with AGO2 and the target that allow these Argonautes to regulate one way over the other. Our kinetic measurements with cleaved products and pairing past the seed support this notion that Argonaute can ‘sense’ the type of target it is bound to. ‘Professional’ slicers may also have similar sets of target interactions that take place past the seed that help favor perfect targets and disfavor partially matched ones. The overall theme for Argonaute is to use base pairing to drive target identification/interaction thus creating an overall affinity for the target — an affinity that is linked to its evolved cellular function.

Conserved properties among RNA-guided endonucleases

In the past few years there has been a lot of excitement and research into a newly discovered mechanism of nucleic acid regulation that uses CRISPR (Clustered regularly interspaced short palindromic repeats) small RNAs (crRNAs, CRISPR RNAs) (Pennisi, 2013; Jiang and Marraffini, 2015). A family of three types of RNA-guided endonucleases known as Cas (CRISPR associated)

proteins bind to crRNAs and use them as guides to cleave complementary sequences. The Cas proteins along with their bound crRNAs are a prokaryotic immune system that combats bacteriophage infection by site specifically cleaving DNA (Jiang and Marraffini, 2015). The interest in CRISPR is not just about the discovery of a new for pathway of viral immunity in prokaryotes rather it is applying Cas proteins to cleave and edit the DNA in other organisms (Jinek et al., 2012). The research and therapeutic potential of CRISPR-Cas rivals the discovery of RNAi where instead of inducing temporal changes at the RNA level, CRISPR cleaves at the DNA level and thus creating a permanent change.

The most studied Cas proteins are the type II family, in particular Cas9. Cas9 uses two RNAs to guide it toward its target, the crRNA that guides Cas9 to its target and a tracrRNA (trans-encoded crRNA) that aids in the processing of the crRNA and complex formation with Cas9 (Deltcheva et al., 2011). Adaptations of having two separate RNAs have been made to create a single-guide RNA (sgRNA) transcript that makes it more amenable in genome editing applications (Jinek et al., 2012). Several structural and biochemical studies have been recently reported investigating how Cas9 bound to a sgRNA finds, binds, and cleaves targets (Sternberg et al., 2014; Jiang and Marraffini, 2015; Sternberg et al., 2015; Taylor et al., 2015). Are there similarities to its 'cousin' endonuclease Argonaute? At first glance, there are several major differences that argue Argonaute and Cas9 are nothing alike. For example, the protein

domain architecture is completely different from Argonaute proteins, the Cas sgRNA is much longer, endogenous crRNA biogenesis is much different, and last but not least, Cas targets DNA. Even with so many differences, there are remarkable parallels in the mechanism of finding and interrogating targets that may represent conserved characteristics of all RNA-guided endonucleases.

The Cas9 protein, similar to Argonaute, searches for its target via a three-dimensional search mechanism (Sternberg et al., 2014). A difference since Cas9 targets DNA, which is double-stranded and precludes nucleic-nucleic acid hybridization/interrogation, is that target find consists of two stages. The first stage is Cas9 finding a PAM (protospacer adjacent motif) sequence that contains a characteristic NGG sequence through its PAM interacting domain (Mojica et al., 2009; Sternberg et al., 2014). Cas9 has an affinity for PAM sequences and this increases when the DNA target just contain multiple PAM sites (Sternberg et al., 2014). Similar to Argonaute, Cas9 shows little to no affinity to DNA without a PAM just like Argonaute having little to no affinity for an RNA that does not bear any resemblance of a seed sequence (Figure S3.5). When Cas9 binds to a PAM sequence it is proposed that there is local strand displacement or R-loop formation that allows for interrogation of the target site (Sternberg et al., 2014). The stretch of ssDNA exposed is ‘inspected’ for complementarity to the guide RNA bound to Cas9. The stretch of guide RNA that is proximal to the PAM and performs the initial interrogation of the target is

referred to as the seed, a name that is very appropriate (Semenova et al., 2011; Wiedenheft et al., 2011; Jinek et al., 2012). Recent structural studies of the Cas9 protein revealed that the seed, similar to the seed created by Argonaute, is also in a pre-helical shape that is made by the protein (Figure 4.1 ;Jiang et al., 2015). Furthermore, mismatch analyses in cell culture reporter assays for the tolerance of mismatches between the sgRNA and target showed little tolerance of single or dinucleotide mismatches in any region of the seed whereas mismatches outside the seed have a lower impact on genome editing (Hsu et al., 2013; Qi et al., 2013). This result is reminiscent of our kinetic measurements that show the seed acts like a cooperative unit, requiring complete pairing to create a stable RISC-target complex.

Figure 4.2

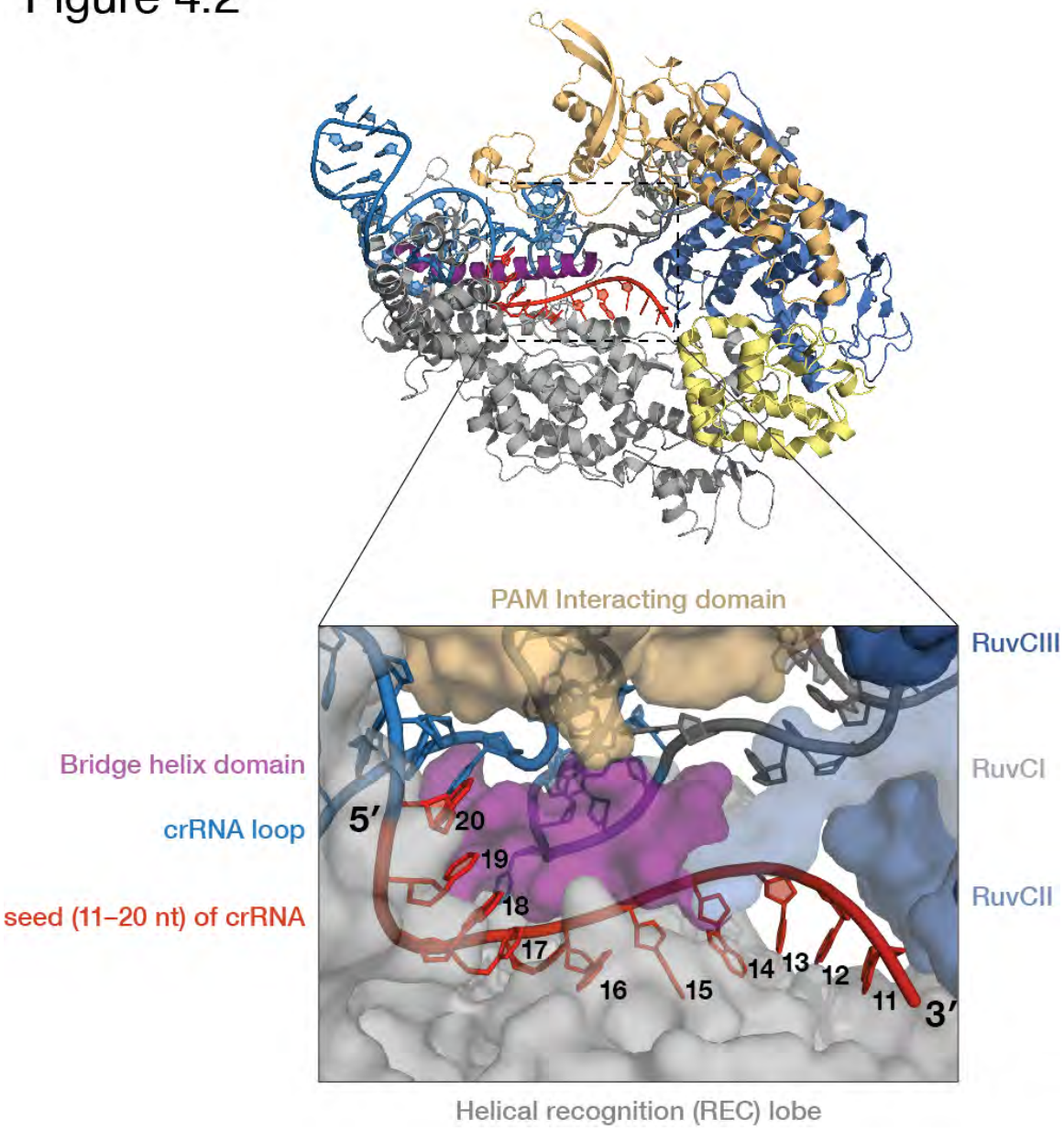

PAM Interacting domain
RuvCIII
Bridge helix domain
RuvCI
crRNA loop
RuvCII
seed (11–20 nt) of crRNA
5′
20
19
18
17
16
15
14
13
12
11
3′
Helical recognition (REC) lobe

Figure 4.2: Pre-organized seed of sgRNA bound to *Streptococcus pyogenes* Cas9

The sgRNA bound to Cas9 changes the shape of the bound RNA to form a pre-helix, similar to Argonaute and its small RNA/DNA. The pre-helical shape is A-form in structure and spans nucleotides 11–20 counting starting from the 3′ end of the sgRNA. Additionally, the REC lobe creates a kink near g15 and g16 that is reminiscent of α-helix 7 of the L2 linker of Argonaute. This kink may confer functions in target site interrogation and binding fidelity.
(Crystal structure solved by: Jiang et al., 2015)

Does the seed of the sgRNA created by Cas9 contain sub-domains? Elegant single-molecule studies by Sternberg, Redding and colleagues evaluated the tolerance for mismatches on Cas9-sgRNA complex binding to a DNA target to understand the mechanism of interrogation (Sternberg et al., 2014). Remarkably, Cas9 shows similar dwell times in regards to seed mismatches that take place at the first two nucleotides where the time bound to the target is ~1000-fold shorter than a fully complementary target even though there is eighteen contiguous base pairs past the mismatches (Sternberg et al., 2014). Furthermore, partial seed matches show the same effect on dwell time (Sternberg et al., 2014), a flattened effect, just like the observed effect 5′ sub-domain seed mismatches have with AGO2-RISC and its RNA target (Figure S3.7). Future studies that systematically look at Cas9-sgRNA complex with seed mismatches to the target may reveal, similar to AGO2, that there is a cooperative effect of the seed that requires ordered and complete base pairing to establish a stable complex. Moreover, relating 'professional' slicer Argonautes dependence on highly complementary targets may reveal that similar to Cas9, there is a certain minimum requirement for contiguous base pairing past the seed and there is a transition point for stable Cas9-sgRNA-target complex formation.

Conclusions

This thesis takes a detailed and quantitative look at how Argonaute proteins find, bind, and cleave their targets. Our studies were some of the first to use single-molecule TIRF microscopy in the field of small RNAs to study miRNA and RNAi pathways. We directly measured and showed how a protein can change biophysical properties of another macromolecule, nucleic acid, so that it is better suited for its cellular function. Furthermore, the imposed changes to the protein bound small RNA guide were more than originally predicted where we have discovered sub-domains in the previously defined functional domains of the guide (Wee et al., 2012) that can be stabilizing and/or de-stabilizing. The powerful single-molecule technique when applied to enzymology allowed us to take measurements that were not possible with ensemble techniques (*e.g.,* monitor product release order). Direct visualization of a photostable RNA target allowed us to uncover that AGO-RISC, like most enzymes, has a higher affinity to substrates than it does to products.

These studies lay a foundation for future studies into understanding the kinetic details of other Argonaute proteins that may have different properties that are best suited for their cellular role. An understanding of the kinetics of slicer Argonautes vs. non-slicing Argonautes may help shed light on some of the true functions for Argonautes, especially with PIWI Argonautes that currently lack extracts that can reconstitute activity. Furthermore, extending our findings

to understand other RNA-guided endonucleases such as Cas proteins can serve to understand a shared enzymology with Argonautes.

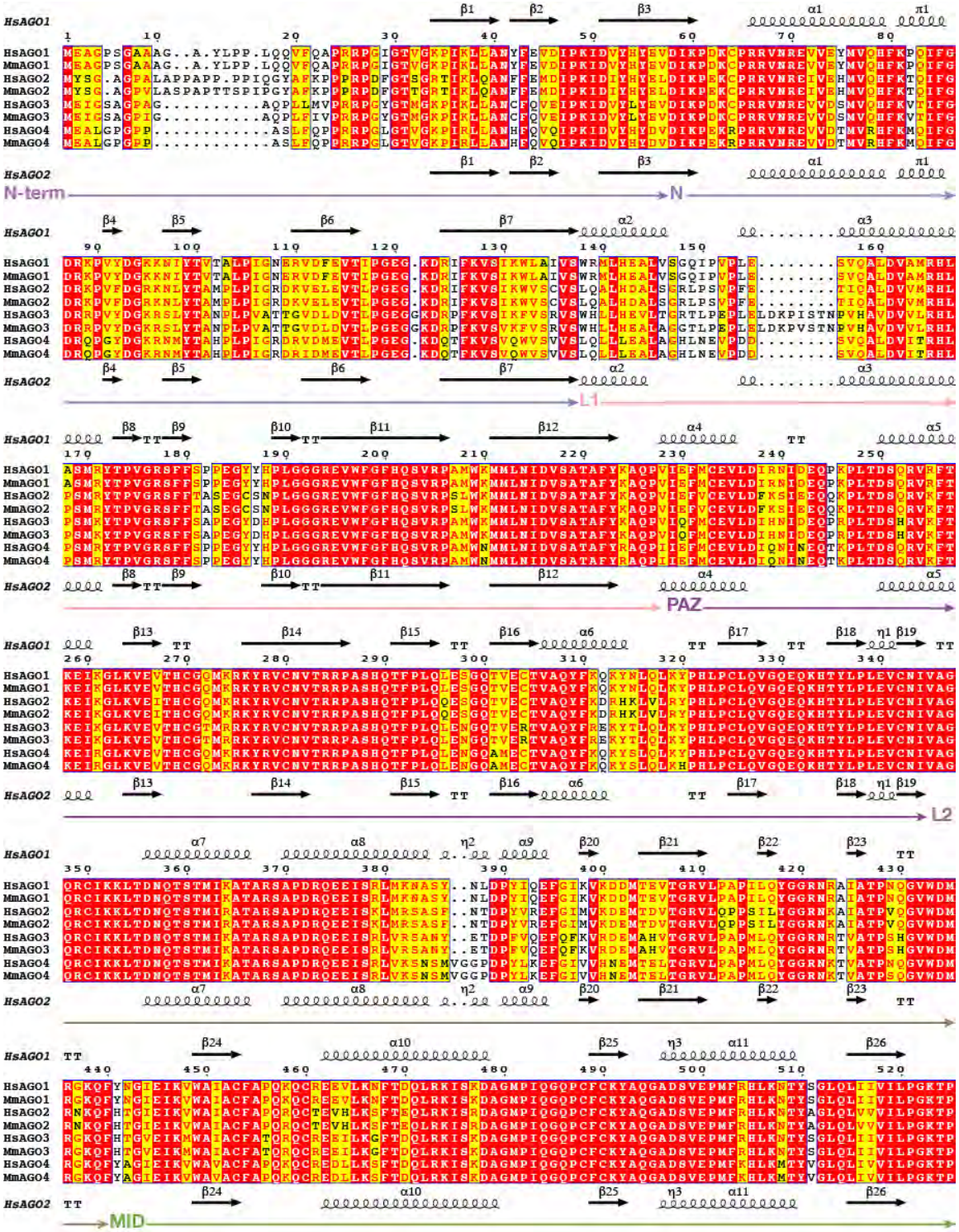
The extension of these studies in the future will help in the development of some of the most promising new therapeutic approaches to treat human diseases. Currently, siRNA-based drugs are being developed and tested in clinical trials and there are several plans to take the CRISPR/Cas9 genome editing technology in the clinic within the next few years. There are and will be several challenges to the use RNA-guided endonucleases as a therapy but mechanistic studies like those presented in this thesis lay a foundation of knowledge that can be built upon for more effective and safer approaches with the use of this remarkable biotechnology.

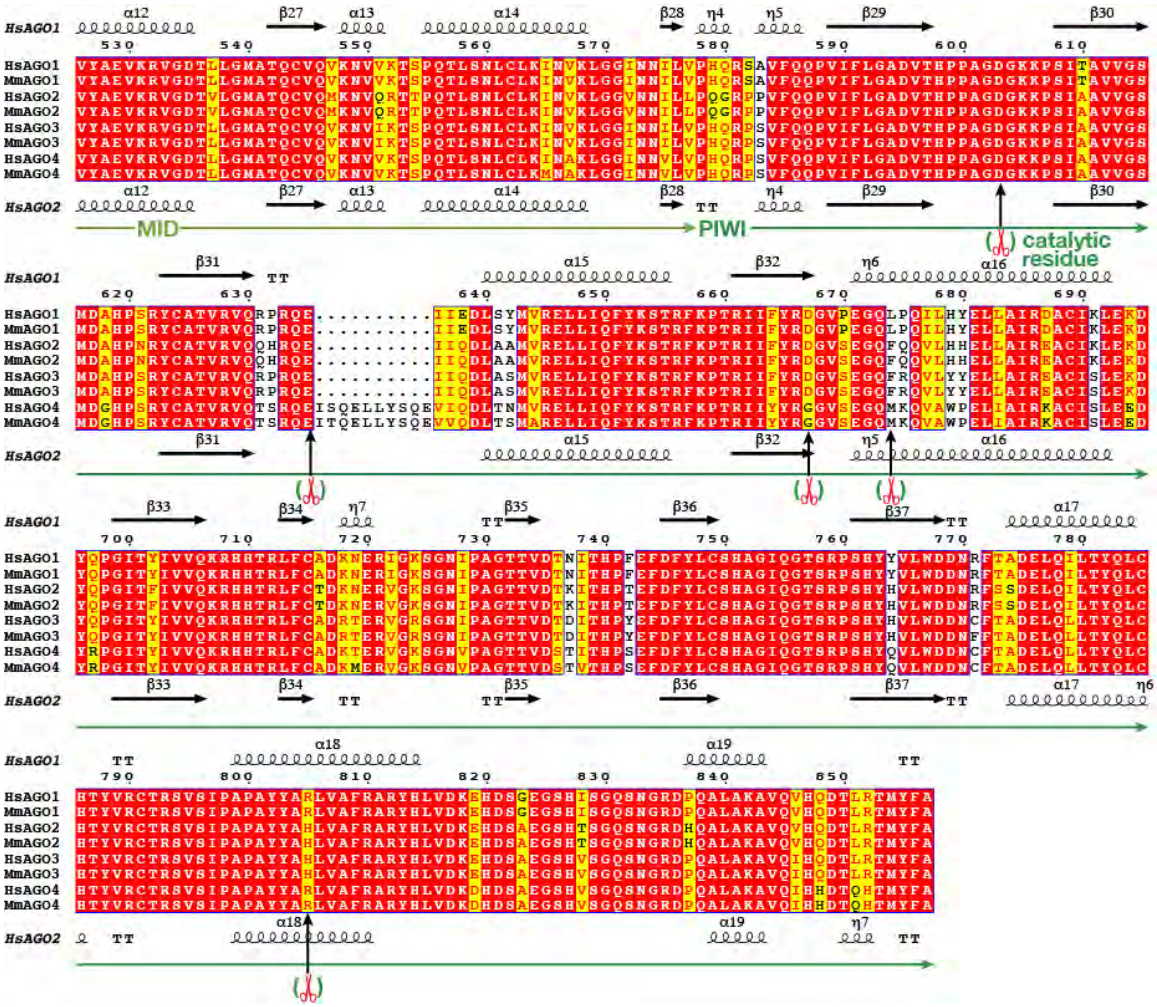
Appendix A: Argonaute Sequence Alignments

A1. Alignment of Human and Mouse AGO family proteins

A2. Alignment of Human and Mouse AGO2 proteins

A1. Protein alignment of mouse and human AGO proteins




MID
PIWI
catalytic residue
HsAGO1
HsAGO2
MmAGO1
MmAGO2
HsAGO3
MmAGO3
HsAGO4
MmAGO4

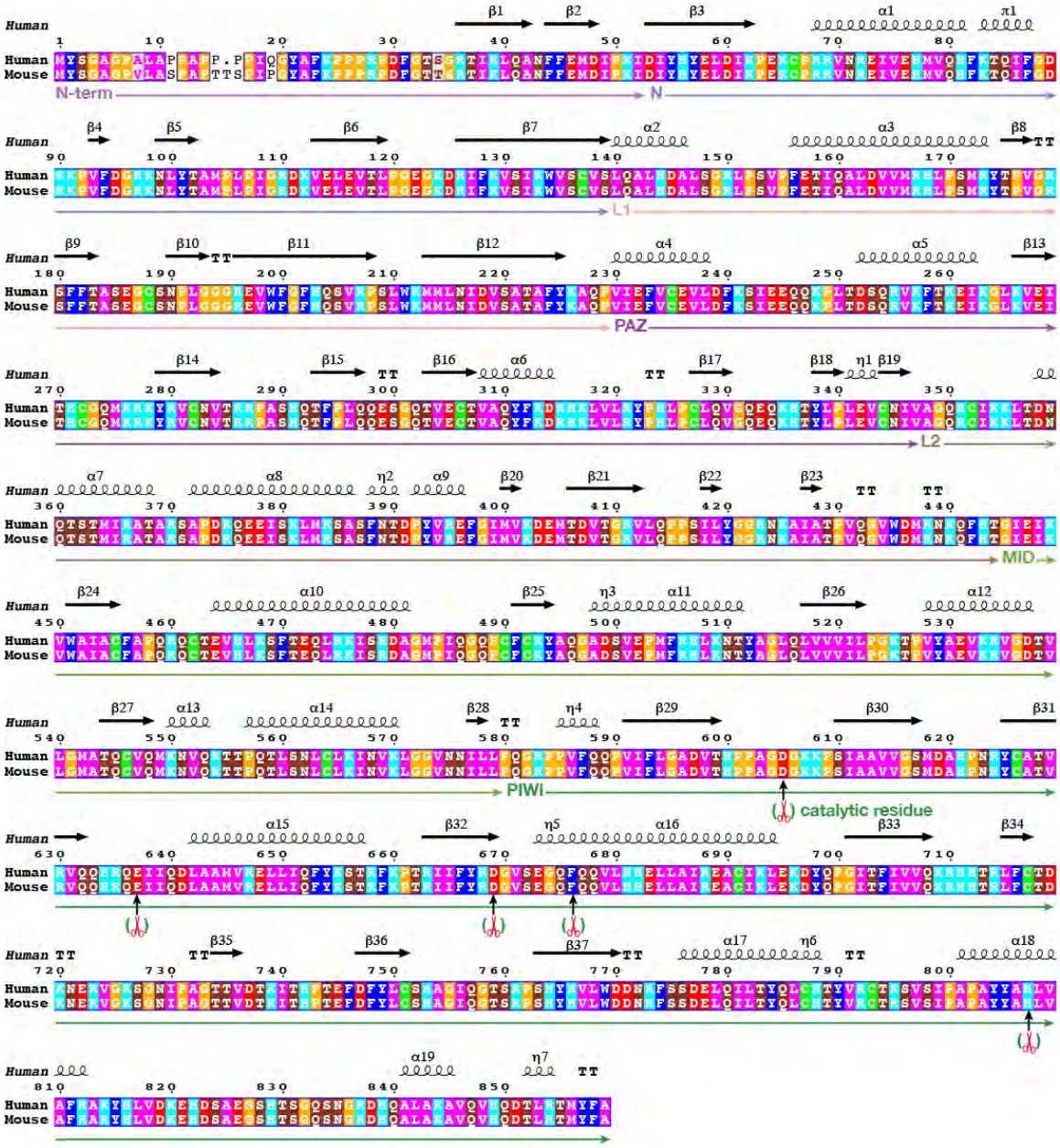
Alignment of human and mouse AGO proteins

Human AGO1 (NP_036331.1), AGO2 (NP_036286.2), AGO3 (NP_079128.2) and AGO4 (NP_060099.2) were aligned along with mouse AGO1 (NP_700452.2), AGO2 (NP_694818.3), AGO3 (NP_700451.2), and AGO4 (NP_694817.2) using T-COFFEE Expresso (BLAST using PDB database secondary structure ;Armougom et al., 2006) and figure was generated using ENDscript to highlight key structural features (Robert and Gouet, 2014). Domains shown below alignment along with scissors to represent catalytic tetrad residues along with a conserved Phe in PIWI required for catalytic activity (Faehnle et al., 2013). Secondary structure shown is derived from PDB structure human AGO1 (above, 4KRF ;Faehnle et al., 2013) and human AGO2 (below, 4W5N ;Schirle et al., 2014). Red = common identity Yellow= some similarities, White = no identity. Differences are slight between human and mice, highlights typically are a difference between 2 AGO proteins and not species.

Amino Acid code:

Single letter code	Three letter code	Amino Acid
I	Iso	Isoleucine
L	Leu	Leucine
V	Val	Valine
F	Phe	Phenylalanine
M	Met	Methionine
C	Cys	Cysteine
A	Ala	Alanine
G	Gly	Glycine
P	Pro	Proline
T	Thr	Threonine
S	Ser	Serine
Y	Tyr	Tyrosine
W	Typ	Tryptophan
Q	Glu	Glutamine
N	Asn	Asparagine
H	His	Histidine
E	Glu	Glutamic acid
D	Asp	Aspartic acid
K	Lys	Lysine
R	Arg	Arginine

A2. Protein alignment of Human and Mouse AGO2



Alignment of human and mouse AGO2

Human AGO2 (NP_036286.2) and mouse AGO2 (NP_694818.3) were aligned using T-COFFEE Expresso (BLAST using PDB database secondary structure ;Armougom et al., 2006) and figure was generated ENDscript to highlight key structural features (Robert and Gouet, 2014). Domains shown below alignment and scissors represent catalytic tetrad residues along with a conserved Phe in PIWI required for catalytic activity (Faehnle et al., 2013). Secondary structure is derived from PDB structure 4W5N (Schirle et al., 2014). White represents identity difference or gap between human and mouse AGO2. Gray represents amino acid difference has similar properties. Other colors represent physicochemical properties of the amino acids: Cyan = polar and positive, Red = polar and negative, Maroon = polar and neutral, Pink = non-polar and aliphatic, Blue = non-polar and aromatic, Orange = proline and glycine, Green = cysteine.

BIBLIOGRAPHY

Aitken, C. E., Marshall, R. A., and Puglisi, J. D. (2008). An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. Biophys J *94*, 1826-1835.

Ambros, V. (2001). Development. Dicing up RNAs. Science *293*, 811-813.

Ambros, V., Bartel, B., Bartel, D. P., Burge, C. B., Carrington, J. C., Chen, X., Dreyfuss, G., Eddy, S. R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., and Tuschl, T. (2003). A uniform system for microRNA annotation. RNA *9*, 277-279.

Ameres, S. L., Martinez, J., and Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. Cell *130*, 101-112.

Ameres, S. L., and Zamore, P. D. (2013). Diversifying microRNA sequence and function. Nat Rev Mol Cell Biol *14*, 475-488.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M. J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J. J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., and Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin, A. A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K. F., Bestor, T., and Hannon, G. J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell *31*, 785-799.

Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G. J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744-747.

Aravin, A. A., Naumova, N. M., Tulin, A. V., Vagin, V. V., Rozovsky, Y. M., and Gvozdev, V. A. (2001). Double-stranded RNA-mediated silencing of genomic

tandem repeats and transposable elements in the *D. melanogaster* germline. Curr Biol *11*, 1017-1027.

Ariyoshi, M., Vassylyev, D. G., Iwasaki, H., Nakamura, H., Shinagawa, H., and Morikawa, K. (1994). Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from E. coli. Cell *78*, 1063-1072.

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic Acids Res *34*, W604-W608.

Azuma-Mukai, A., Oguri, H., Mituyama, T., Qian, Z. R., Asai, K., Siomi, H., and Siomi, M. C. (2008). Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. Proc Natl Acad Sci U S A *105*, 7964-7969.

Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P., and Bartel, D. P. (2008). The impact of microRNAs on protein output. Nature *455*, 64-71.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., Chaves, D. A., Gu, W., Vasale, J. J., Duan, S., Conte, D. J., Luo, S., Schroth, G. P., Carrington, J. C., Bartel, D. P., and Mello, C. C. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. Mol Cell *31*, 67-78.

Bazzini, A. A., Lee, M. T., and Giraldez, A. J. (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. Science *336*, 233-237.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde,

E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev *20*, 1885-1898.

Beitzinger, M., Peters, L., Zhu, J. Y., Kremmer, E., and Meister, G. (2007). Identification of human microRNA targets from isolated argonaute protein complexes. RNA Biol *4*, 76-84.

Berg, O. G., and von Hippel, P. H. (1985). Diffusion-controlled macromolecular interactions. Annu Rev Biophys Biophys Chem *14*, 131-160.

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Bernstein, E., Kim, S. Y., Carmell, M. A., Murchison, E. P., Alcorn, H., Li, M. Z., Mills, A. A., Elledge, S. J., Anderson, K. V., and Hannon, G. J. (2003). Dicer is essential for mouse development. Nat Genet *35*, 215-217.

Bernstein, P., Peltz, S. W., and Ross, J. (1989). The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. Mol Cell Biol *9*, 659-670.

Betancur, J. G., and Tomari, Y. (2012). Dicer is dispensable for asymmetric RISC loading in mammals. RNA *18*, 24-30.

Bohnsack, M. T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a Ran GTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA *10*, 185-191.

Boland, A., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2011). Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. Proc Natl Acad Sci U S A *108*, 10466-10471.

Boland, A., Tritschler, F., Heimstadt, S., Izaurralde, E., and Weichenrieder, O. (2010). Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. EMBO Rep *11*, 522-527.

Braun, J. E., Huntzinger, E., Fauser, M., and Izaurralde, E. (2011). GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. Mol Cell *44*, 120-133.

Braun, J. E., Truffault, V., Boland, A., Huntzinger, E., Chang, C. T., Haas, G., Weichenrieder, O., Coles, M., and Izaurralde, E. (2012). A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation. Nat Struct Mol Biol *19*, 1324-1331.

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G. J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell *128*, 1089-1103.

Brennecke, J., Stark, A., Russell, R. B., and Cohen, S. M. (2005). Principles of microRNA-target recognition. PLoS Biol *3*, e85.

Broderick, J. A., Salomon, W. E., Ryder, S. P., Aronin, N., and Zamore, P. D. (2011). Argonaute protein identity and pairing geometry determine cooperativity in mammalian RNA silencing. RNA *17*, 1858-1869.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., and Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. Science *320*, 1185-1190.

Buhler, M., Verdel, A., and Moazed, D. (2006). Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. Cell *125*, 873-886.

Cava, F., Hidalgo, A., and Berenguer, J. (2009). *Thermus thermophilus* as biological model. Extremophiles *13*, 213-231.

Cenik, E. S., and Zamore, P. D. (2011). Argonaute proteins. Curr Biol *21*, R446-R449.

Cerutti, L., Mian, N., and Bateman, A. (2000). Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi

domain. Trends in Biochemical Sciences 481-482.

Chandradoss, S. D., Schirle, N. T., Szczepaniak, M., MacRae, I. J., and Joo, C. (2015). A Dynamic Search Process Underlies MicroRNA Targeting. Cell *162*, 96-107.

Chang, C. T., Hain, T. C., Hutton, J. R., and Wetmur, J. G. (1974). Effects of microscopic and macroscopic viscosity on the rate of renaturation of DNA. Biopolymers *13*, 1847-1858.

Chapados, B. R., Chai, Q., Hosfield, D. J., Qiu, J., Shen, B., and Tainer, J. A. (2001). Structural biochemistry of a type 2 RNase H: RNA primer recognition and removal during DNA replication. J Mol Biol *307*, 541-556.

Chekulaeva, M., Mathys, H., Zipprich, J. T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. Nat Struct Mol Biol *18*, 1218-1226.

Cheloufi, S., Dos Santos, C. O., Chong, M. M., and Hannon, G. J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature *465*, 584-589.

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.

Chiang, H. R., Schoenfeld, L. W., Ruby, J. G., Auyeung, V. C., Spies, N., Baek, D., Johnston, W. K., Russ, C., Luo, S., Babiarz, J. E., Blelloch, R., Schroth, G. P., Nusbaum, C., and Bartel, D. P. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev *24*, 992-1009.

Chung, W. J., Okamura, K., Martin, R., and Lai, E. C. (2008). Endogenous RNA Interference Provides a Somatic Defense against *Drosophila* Transposons.

Curr Biol *18*, 795-802.

Church, G. M. (1984). Genomic Sequencing. Proceedings of the National Academy of Sciences *81*, 1991-1995.

Cora, E., Pandey, R. R., Xiol, J., Taylor, J., Sachidanandam, R., McCarthy, A. A., and Pillai, R. S. (2014). The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs. RNA

Crawford, D. J., Hoskins, A. A., Friedman, L. J., Gelles, J., and Moore, M. J. (2008). Visualizing the splicing of single pre-mRNA molecules in whole cell extract. RNA *14*, 170-179.

Crocker, J. C., and Grier, D. G. (1996). Methods of Digital Video Microscopy for Colloidal Studies. Journal of Colloid and Interface Science *179*, 298-310.

Czech, B., Malone, C. D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J. A., Sachidanandam, R., Hannon, G. J., and Brennecke, J. (2008). An endogenous small interfering RNA pathway in *Drosophila*. Nature *453*, 798-802.

Czech, B., Zhou, R., Erlich, Y., Brennecke, J., Binari, R., Villalta, C., Gordon, A., Perrimon, N., and Hannon, G. J. (2009). Hierarchical rules for Argonaute loading in *Drosophila*. Mol Cell *36*, 445-456.

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., Buhecha, H. R., Gilchrist, M. J., Howe, K. L., Stark, R., Matthews, N., Berezikov, E., Ketting, R. F., Tavare, S., and Miska, E. A. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. Mol Cell *31*, 79-90.

Dave, R., Terry, D. S., Munro, J. B., and Blanchard, S. C. (2009). Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging. Biophys J *96*, 2371-2381.

Davies, D. R., Goryshin, I. Y., Reznikoff, W. S., and Rayment, I. (2000). Three-

dimensional structure of the Tn5 synaptic complex transposition intermediate. Science *289*, 77-85.

De Fazio, S., Bartonicek, N., Di Giacomo, M., Abreu-Goodger, C., Sankar, A., Funaya, C., Antony, C., Moreira, P. N., Enright, A. J., and O'Carroll, D. (2011). The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature *480*, 259-263.

De, N., Young, L., Lau, P. W., Meisner, N. C., Morrissey, D. V., and MacRae, I. J. (2013). Highly complementary target RNAs promote release of guide RNAs from human Argonaute2. Mol Cell *50*, 344-355.

Deerberg, A., Willkomm, S., and Restle, T. (2013). Minimal mechanistic model of siRNA-dependent target RNA slicing by recombinant human Argonaute 2 protein. Proc Natl Acad Sci U S A *110*, 17850-17855.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602-607.

Denzler, R., Agarwal, V., Stefano, J., Bartel, D. P., and Stoffel, M. (2014). Assessing the ceRNA Hypothesis with Quantitative Measurements of miRNA and Target Abundance. Mol Cell *54*, 766-776.

Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res *11*, 1475-1489.

Ding, H., Schwarz, D. S., Keene, A., Affar el, B., Fenton, L., Xia, X., Shi, Y., Zamore, P. D., and Xu, Z. (2003). Selective silencing by RNAi of a dominant allele that causes amyotrophic lateral sclerosis. Aging Cell *2*, 209-217.

Dorsett, Y., and Tuschl, T. (2004). siRNAs: applications in functional genomics and potential as therapeutics. Nat Rev Drug Discov *3*, 318-329.

Egli, M., and Saenger, W. (1988). Principles of Nucleic Acid Structure (Springer

Advanced Texts in Chemistry) Springer).

Eichhorn, S. W., Guo, H., McGeary, S. E., Rodriguez-Mias, R. A., Shin, C., Baek, D., Hsu, S. H., Ghoshal, K., Villen, J., and Bartel, D. P. (2014). mRNA Destabilization Is the Dominant Effect of Mammalian MicroRNAs by the Time Substantial Repression Ensues. Mol Cell *56*, 104-115.

Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001a). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev *15*, 188-200.

Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001b). Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. EMBO J *20*, 6877-6888.

Elkayam, E., Kuhn, C. D., Tocilj, A., Haase, A. D., Greene, E. M., Hannon, G. J., and Joshua-Tor, L. (2012). The Structure of Human Argonaute-2 in Complex with miR-20a. Cell *150*, 100-110.

Elmen, J., Lindow, M., Schutz, S., Lawrence, M., Petri, A., Obad, S., Lindholm, M., Hedtjarn, M., Hansen, H. F., Berger, U., Gullans, S., Kearney, P., Sarnow, P., Straarup, E. M., and Kauppinen, S. (2008). LNA-mediated microRNA silencing in non-human primates. Nature *452*, 896-899.

Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008). GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. Nat Struct Mol Biol *15*, 346-353.

Eulalio, A., Tritschler, F., and Izaurralde, E. (2009). The GW182 protein family in animal cells: new insights into domains required for miRNA-mediated gene silencing. RNA *15*, 1433-1442.

Fabian, M. R., Cieplak, M. K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T. F., and Sonenberg, N. (2011). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. Nat Struct Mol Biol *18*, 1211-1217.

Faehnle, C. R., Elkayam, E., Haase, A. D., Hannon, G. J., and Joshua-Tor, L. (2013). The Making of a Slicer: Activation of Human Argonaute-1. Cell Rep *3*, 1901-1909.

Fagard, M., Boutet, S., Morel, J.-B., Bellini, C., and Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc Natl Acad Sci U S A *97*, 11650-11654.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature *391*, 806-811.

Flores-Jasso, C. F., Salomon, W. E., and Zamore, P. D. (2013). Rapid and specific purification of Argonaute-small RNA complexes from crude cell lysates. RNA *19*, 271-279.

Forstemann, K., Horwich, M. D., Wee, L., Tomari, Y., and Zamore, P. D. (2007). *Drosophila* microRNAs are sorted into functionally distinct Argonaute complexes after production by Dicer-1. Cell *130*, 287-297.

Forstemann, K., Tomari, Y., Du, T., Vagin, V. V., Denli, A. M., Bratu, D. P., Klattenhoff, C., Theurkauf, W. E., and Zamore, P. D. (2005). Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol *3*, e236.

Frank, F., Sonenberg, N., and Nagar, B. (2010). Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. Nature *465*, 818-822.

Freier, S. M., and Altmann, K. H. (1997). The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA:RNA duplexes. Nucleic Acids Res *25*, 4429-4443.

Friedman, L. J., Chung, J., and Gelles, J. (2006). Viewing dynamic assembly of

molecular complexes by multi-wavelength single-molecule fluorescence. Biophys J *91*, 1023-1031.

Friedman, L. J., and Gelles, J. (2015). Multi-wavelength single-molecule fluorescence analysis of transcription mechanisms. Methods

Friedman, L. J., Mumm, J. P., and Gelles, J. (2013). RNA polymerase approaches its promoter without long-range sliding along DNA. Proc Natl Acad Sci U S A *110*, 9740-9745.

Friedman, R. C., Farh, K. K., Burge, C. B., and Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome Res *19*, 92-105.

Galau, G. A., Britten, R. J., and Davidson, E. H. (1977). Studies on nucleic acid reassociation kinetics: rate of hybridization of excess RNA with DNA, compared to the rate of DNA renaturation. Proc Natl Acad Sci U S A *74*, 1020-1023.

Gerbasi, V. R., Golden, D. E., Hurtado, S. B., and Sontheimer, E. J. (2010). Proteomic identification of *Drosophila* siRNA-associated factors. Mol Cell Proteomics

Ghildiyal, M., Seitz, H., Horwich, M. D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E. L., Zapp, M. L., Weng, Z., and Zamore, P. D. (2008). Endogenous siRNAs Derived from Transposons and mRNAs in *Drosophila* Somatic Cells. Science *320*, 1077-1081.

Ghildiyal, M., Xu, J., Seitz, H., Weng, Z., and Zamore, P. D. (2010). Sorting of *Drosophila* small silencing RNAs partitions microRNA* strands into the RNA interference pathway. RNA *16*, 43-56.

Ghildiyal, M., and Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. Nat Rev Genet *10*, 94-108.

Girard, A., Sachidanandam, R., Hannon, G. J., and Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins.

Nature *442*, 199-202.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res *38*, W695-W699.

Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Grimson, A., Farh, K. K., Johnston, W. K., Garrett-Engele, P., Lim, L. P., and Bartel, D. P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell *27*, 91-105.

Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B. J., Chiang, H. R., King, N., Degnan, B. M., Rokhsar, D. S., and Bartel, D. P. (2008). Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature *455*, 1193-1197.

Grivna, S. T., Beyret, E., Wang, Z., and Lin, H. (2006a). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev *20*, 1709-1714.

Grivna, S. T., Pyhtila, B., and Lin, H. (2006b). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc Natl Acad Sci U S A *103*, 13415-13420.

Gryaznov, S., and Schultz, R. G. (1994). Stabilization of DNA:DNA and DNA/RNA duplexes by substitution of 2'-deoxyadenosine with 2'-deoxy-2-aminoadenosine. Tetrahedron Letters *35*, 2489-2492.

Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M. C. (2007). A Slicer-Mediated Mechanism for Repeat-Associated siRNA 5′ End Formation *in Drosophi*la. Science *315*, 1587-1590.

Guo, H., Ingolia, N. T., Weissman, J. S., and Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*,

835-840.

Gurtan, A. M., and Sharp, P. A. (2013). The role of miRNAs in regulating gene expression networks. J Mol Biol *425*, 3582-3600.

Ha, M., and Kim, V. N. (2014). Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol *15*, 509-524.

Haase, A. D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep *6*, 961-967.

Haley, B., Tang, G., and Zamore, P. D. (2003). In vitro analysis of RNA interference in *Drosophila melanogaster*. Methods *30*, 330-336.

Haley, B., and Zamore, P. D. (2004). Kinetic analysis of the RNAi enzyme complex. Nat Struct Mol Biol *11*, 599-606.

Hamilton, A. J., and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. Science *286*, 950-952.

Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature *404*, 293-296.

Han, J., Lee, Y., Yeom, K. H., Kim, Y. K., Jin, H., and Kim, V. N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev *18*, 3016-3027.

Hauptmann, J., Dueck, A., Harlander, S., Pfaff, J., Merkl, R., and Meister, G. (2013). Turning catalytically inactive human Argonaute proteins into active slicer enzymes. Nat Struct Mol Biol *20*, 814-817.

Hendrickson, D. G., Hogan, D. J., McCullough, H. L., Myers, J. W., Herschlag, D., Ferrell, J. E., and Brown, P. O. (2009). Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. PLoS Biol *7*, e1000238.

Herschlag, D. (1991). Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules in vivo: more isn't always better. Proc Natl Acad Sci U S A *88*, 6921-6925.

Herzog, V. A., and Ameres, S. L. (2015). Approaching the Golden Fleece a Molecule at a Time: Biophysical Insights into Argonaute-Instructed Nucleic Acid Interactions. Mol Cell *59*, 4-7.

Hock, J., and Meister, G. (2008). The Argonaute protein family. Genome Biol *9*, 210.

Horwich, M. D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P. D. (2007). The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr Biol *17*, 1265-1272.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G., and Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol *31*, 827-832.

Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet *12*, 99-110.

Hutton, J. R., and Wetmur, J. G. (1973). Renaturation of bacteriophage phiX174 DNA-RNA hybrid: RNA length effect and nucleation rate constant. J Mol Biol *77*, 495-500.

Hutvágner, G., McLachlan, J., Pasquinelli, A. E., Balint, É., Tuschl, T., and Zamore, P. D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the *let-7* small temporal RNA. Science *293*, 834-838.

Hutvagner, G., and Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol *9*, 22-32.

Hutvagner, G., Simard, M. J., Mello, C. C., and Zamore, P. D. (2004). Sequence-specific inhibition of small RNA function. PLoS Biol *2*, E98.

Hutvagner, G., and Zamore, P. D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science *297*, 2056-2060.

Ikeda, K., Satoh, M., Pauley, K. M., Fritzler, M. J., Reeves, W. H., and Chan, E. K. (2006). Detection of the argonaute protein Ago2 and microRNAs in the RNA induced silencing complex (RISC) using a monoclonal antibody. J Immunol Methods *317*, 38-44.

Iki, T., Yoshikawa, M., Nishikiori, M., Jaudal, M. C., Matsumoto-Yokoyama, E., Mitsuhara, I., Meshi, T., and Ishikawa, M. (2010). In Vitro Assembly of Plant RNA-Induced Silencing Complexes Facilitated by Molecular Chaperone HSP90. Mol Cell *39*, 282-291.

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., and Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218-223.

Ipsaro, J. J., Haase, A. D., Knott, S. R., Joshua-Tor, L., and Hannon, G. J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature *491*, 279-283.

Iwakawa, H. O., and Tomari, Y. (2013). Molecular insights into microRNA-mediated translational repression in plants. Mol Cell *52*, 591-601.

Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T., and Tomari, Y. (2010). Hsc70/Hsp90 Chaperone Machinery Mediates ATP-Dependent RISC Loading of Small RNA Duplexes. Mol Cell *39*, 292-299.

Iwasaki, S., Sasaki, H. M., Sakaguchi, Y., Suzuki, T., Tadakuma, H., and Tomari, Y. (2015a). Defining fundamental steps in the assembly of the Drosophila RNAi enzyme complex. Nature *521*, 533-536.

Iwasaki, Y. W., Siomi, M. C., and Siomi, H. (2015b). PIWI-Interacting RNA: Its Biogenesis and Functions. Annu Rev Biochem *84*, 405-433.

Jiang, F., Ye, X., Liu, X., Fincher, L., McKearin, D., and Liu, Q. (2005). Dicer-1 and R3D1-L catalyze microRNA maturation in *Drosophila*. Genes Dev *19*, 1674-1679.

Jiang, F., Zhou, K., Ma, L., Gressel, S., and Doudna, J. A. (2015). STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. Science *348*, 1477-1481.

Jiang, W., and Marraffini, L. A. (2015). CRISPR-Cas: New Tools for Genetic Manipulations from Bacterial Immunity Systems. Annu Rev Microbiol *69*, 209-228.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.

Johnston, M., Geoffroy, M. C., Sobala, A., Hay, R., and Hutvagner, G. (2010). HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. Mol Biol Cell *21*, 1462-1469.

Jonas, S., and Izaurralde, E. (2013). The role of disordered protein regions in the assembly of decapping complexes and RNP granules. Genes Dev *27*, 2628-2641.

Jonas, S., and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet *16*, 421-433.

Joo, C., and Ha, T. (2012). Preparing sample chambers for single-molecule FRET. Cold Spring Harb Protoc *2012*, 1104-1108.

Jung, S. R., Kim, E., Hwang, W., Shin, S., Song, J. J., and Hohng, S. (2013). Dynamic anchoring of the 3'-end of the guide strand controls the target dissociation of Argonaute-guide complex. J Am Chem Soc *135*, 16865-16871.

Kamminga, L. M., Luteijn, M. J., den Broeder, M. J., Redl, S., Kaaij, L. J., Roovers, E. F., Ladurner, P., Berezikov, E., and Ketting, R. F. (2010). Hen1

is required for oocyte development and piRNA stability in zebrafish. EMBO J *29*, 3688-3700.

Katayanagi, K., Okumura, M., and Morikawa, K. (1993). Crystal structure of Escherichia coli RNase HI in complex with Mg2+ at 2.8 A resolution: proof for a single Mg(2+)-binding site. Proteins *17*, 337-346.

Kawamata, T., Seitz, H., and Tomari, Y. (2009). Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. Nat Struct Mol Biol *16*, 953-960.

Kawamata, T., and Tomari, Y. (2010). Making RISC. Trends Biochem Sci *35*, 368-376.

Kawamata, T., Yoda, M., and Tomari, Y. (2011). Multilayer checkpoints for microRNA authenticity during RISC assembly. EMBO Rep *12*, 944-949.

Kawaoka, S., Izumi, N., Katsuma, S., and Tomari, Y. (2011). 3′ end formation of PIWI-interacting RNAs in vitro. Mol Cell *43*, 1015-1022.

Khvorova, A., Reynolds, A., and Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell *115*, 209-216.

Kim, V. N., Han, J., and Siomi, M. C. (2009). Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol *10*, 126-139.

Kobayashi, H., and Tomari, Y. (2015). RISC assembly: Coordination between small RNAs and Argonaute proteins. Biochim Biophys Acta

Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., and Rajewsky, N. (2005). Combinatorial microRNA target predictions. Nat Genet *37*, 495-500.

Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., and Stoffel, M. (2005). Silencing of microRNAs in vivo with 'antagomirs'. Nature *438*, 685-689.

Kuhn, C. D., and Joshua-Tor, L. (2013). Eukaryotic Argonautes come into focus.

Trends Biochem Sci *38*, 263-271.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T. W., Hata, K., Li, E., Matsuda, Y., Kimura, T., Okabe, M., Sakaki, Y., Sasaki, H., and Nakano, T. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev *22*, 908-917.

Kwak, P. B., and Tomari, Y. (2012). The N domain of Argonaute drives duplex unwinding during RISC assembly. Nat Struct Mol Biol *19*, 145-151.

Lai, E. C. (2002). MicroRNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet *30*, 363-364.

Lai, E. C., and Posakony, J. W. (1998). Regulation of *Drosophila* neurogenesis by RNA:RNA duplexes? Cell *93*, 1103-1104.

Lai, L., Yokota, H., Hung, L. W., Kim, R., and Kim, S. H. (2000). Crystal structure of archaeal RNase HII: a homologue of human major RNase H. Structure *8*, 897-904.

Lanford, R. E., Hildebrandt-Eriksen, E. S., Petri, A., Persson, R., Lindow, M., Munk, M. E., Kauppinen, S., and Orum, H. (2010). Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. Science *327*, 198-201.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science *294*, 858-862.

Lau, N. C., Seto, A. G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.

P., and Kingston, R. E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.

Leaman, D., Chen, P. Y., Fak, J., Yalcin, A., Pearce, M., Unnerstall, U., Marks, D. S., Sander, C., Tuschl, T., and Gaul, U. (2005). Antisense-mediated depletion reveals essential and specific functions of microRNAs in *Drosophila* development. Cell *121*, 1097-1108.

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. Cell *75*, 843-854.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.

Lee, Y., Hur, I., Park, S. Y., Kim, Y. K., Suh, M. R., and Kim, V. N. (2006). The role of PACT in the RNA silencing pathway. EMBO J *25*, 522-532.

Lee, Y., Jeon, K., Lee, J. T., Kim, S., and Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. EMBO J *21*, 4663-4670.

Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., and Kim, V. N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. EMBO J *23*, 4051-4060.

Lee, Y. S., Nakahara, K., Pham, J. W., Kim, K., He, Z., Sontheimer, E. J., and Carthew, R. W. (2004b). Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.

Leuschner, P. J., Ameres, S. L., Kueng, S., and Martinez, J. (2006). Cleavage of the siRNA passenger strand during RISC assembly in human cells. EMBO Rep *7*, 314-320.

Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are

microRNA targets. Cell *120*, 15-20.

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.

Li, L., Yu, C., Gao, H., and Li, Y. (2010). Argonaute proteins: potential biomarkers for human colon cancer. BMC Cancer *10*, 38.

Li, X. Z., Roy, C. K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B. W., Xu, J., Moore, M. J., Schimenti, J. C., Weng, Z., and Zamore, P. D. (2013a). An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes. Mol Cell *50*, 67-81.

Li, X. Z., Roy, C. K., Moore, M. J., and Zamore, P. D. (2013b). Defining piRNA primary transcripts. Cell Cycle *12*, 1657-1658.

Lian, S. L., Li, S., Abadal, G. X., Pauley, B. A., Fritzler, M. J., and Chan, E. K. (2009). The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. RNA *15*, 804-813.

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature *433*, 769-773.

Lima, W. F., Wu, H., Nichols, J. G., Sun, H., Murray, H. M., and Crooke, S. T. (2009). Binding and cleavage specificities of human Argonaute2. J Biol Chem *284*, 26017-26028.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. Nature *426*, 465-469.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004a). NMR assignment of the *Drosophila* Argonaute2 PAZ domain. J Biomol NMR *29*, 421-422.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004b). Nucleic acid 3′-end

recognition by the Argonaute2 PAZ domain. Nat Struct Mol Biol *11*, 576-577.

Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., Hammond, S. M., Joshua-Tor, L., and Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.

Liu, X., Jiang, F., Kalidas, S., Smith, D., and Liu, Q. (2006). Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. RNA *12*, 1514-1520.

Liu, X., Park, J. K., Jiang, F., Liu, Y., McKearin, D., and Liu, Q. (2007). Dicer-1, but not Loquacious, is critical for assembly of miRNA-induced silencing complexes. RNA *13*, 2324-2329.

Liu, Y., Tan, H., Tian, H., Liang, C., Chen, S., and Liu, Q. (2011). Autoantigen La promotes efficient RNAi, antiviral response, and transposon silencing by facilitating multiple-turnover RISC catalysis. Mol Cell *44*, 502-508.

Liu, Y., Ye, X., Jiang, F., Liang, C., Chen, D., Peng, J., Kinch, L. N., Grishin, N. V., and Liu, Q. (2009). C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. Science *325*, 750-753.

Llave, C., Xie, Z., Kasschau, K. D., and Carrington, J. C. (2002). Cleavage of *Scarecrow-Like* mRNA Targets Directed by a Class of *Arabidopsis* miRNA. Science *297*, 2053-2056.

Longfellow, C. E., Kierzek, R., and Turner, D. H. (1990). Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. Biochemistry *29*, 278-285.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.

Ma, J. B., Ye, K., and Patel, D. J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature *429*, 318-322.

Ma, J. B., Yuan, Y. R., Meister, G., Pei, Y., Tuschl, T., and Patel, D. J. (2005).

Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. Nature *434*, 666-670.

MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V., and Doudna, J. A. (2008). In vitro reconstitution of the human RISC-loading complex. Proc Natl Acad Sci U S A *105*, 512-517.

Makarova, K. S., Wolf, Y. I., Snir, S., and Koonin, E. V. (2011). Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. J Bacteriol *193*, 6039-6056.

Mangus, D. A., Evans, M. C., and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol *4*, 223.

Martinez, J., and Tuschl, T. (2004). RISC is a 5′ phosphomonoester-producing RNA endonuclease. Genes Dev *18*, 975-980.

Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., and Tuschl, T. (2002). Single-Stranded Antisense siRNAs Guide Target RNA Cleavage in RNAi. Cell *110*, 563-574.

Matranga, C., Tomari, Y., Shin, C., Bartel, D. P., and Zamore, P. D. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell *123*, 607-620.

Matzke, M. A., and Matzke, A. (1995). How and Why Do Plants Inactivate Homologous (Trans)genes? Plant Physiol *107*, 679-685.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. Nat Rev Genet *14*, 447-459.

Meister, G., Landthaler, M., Dorsett, Y., and Tuschl, T. (2004a). Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. RNA *10*, 544-550.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004b). Human Argonaute2 mediates RNA cleavage targeted by

miRNAs and siRNAs. Mol Cell *15*, 185-197.

Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M. C. (2005). Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. Genes Dev *19*, 2837-2848.

Miyoshi, T., Takeuchi, A., Siomi, H., and Siomi, M. C. (2010). A direct role for Hsp90 in pre-RISC formation in *Drosophila*. Nat Struct Mol Biol *17*, 1024-1026.

Mochizuki, K., Fine, N. A., Fujisawa, T., and Gorovsky, M. A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689-699.

Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology *155*, 733-740.

Moore, M. J., and Query, C. C. (2000). Joining of RNAs by splinted ligation. Methods Enzymol *317*, 109-123.

Moretti, F., Kaiser, C., Zdanowicz-Specht, A., and Hentze, M. W. (2012). PABP and the poly(A) tail augment microRNA repression by facilitated miRISC binding. Nat Struct Mol Biol *19*, 603-608.

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A. K., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of Ribonucleoproteins containing numerous microRNAs. Genes Dev *16*, 720-728.

Nakanishi, K., Ascano, M., Gogakos, T., Ishibe-Murakami, S., Serganov, A. A., Briskin, D., Morozov, P., Tuschl, T., and Patel, D. J. (2013). Eukaryote-specific insertion elements control human ARGONAUTE slicer activity. Cell Rep *3*, 1893-1900.

Nakanishi, K., Weinberg, D. E., Bartel, D. P., and Patel, D. J. (2012). Structure of yeast Argonaute with guide RNA. Nature *486*, 368-374.

Napoli, C., Lemieux, C., and Jorgensen, R. A. (1990). Introduction of a chimeric chalcone synthase gene into Petunia results in reversible co-suppression of homologous genes in trans. Plant Cell *2*, 279-289.

Nelson, J. W., Martin, F. H., and Tinoco, I. J. (1981). DNA and RNA oligomer thermodynamics: the effect of mismatched bases on double-helix stability. Biopolymers *20*, 2509-2531.

Nygaard, A. P., and Hall, B. D. (1964). Formation and properties of RNA-DNA complexes. Journal of Molecular Biology *9*, 125-142.

Nykanen, A., Haley, B., and Zamore, P. D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell *107*, 309-321.

O'Carroll, D., Mecklenbrauker, I., Das, P. P., Santana, A., Koenig, U., Enright, A. J., Miska, E. A., and Tarakhovsky, A. (2007). A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. Genes Dev *21*, 1999-2004.

O'Toole, A. S., Miller, S., and Serra, M. J. (2005). Stability of 3' double nucleotide overhangs that model the 3' ends of siRNA. RNA *11*, 512-516.

Obad, S., dos Santos, C. O., Petri, A., Heidenblad, M., Broom, O., Ruse, C., Fu, C., Lindow, M., Stenvang, J., Straarup, E. M., Hansen, H. F., Koch, T., Pappin, D., Hannon, G. J., and Kauppinen, S. (2011). Silencing of microRNA families by seed-targeting tiny LNAs. Nat Genet *43*, 371-378.

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., and Suzuki, T. (2007). The 3′ termini of mouse Piwi-interacting RNAs are 2′-O-methylated. Nat Struct Mol Biol *14*, 349-350.

Okamura, K., Balla, S., Martin, R., Liu, N., and Lai, E. C. (2008a). Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster*. Nat Struct Mol Biol *15*, 581-590.

Okamura, K., Chung, W. J., Ruby, J. G., Guo, H., Bartel, D. P., and Lai, E. C.

(2008b). The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. Nature *453*, 803-806.

Okamura, K., Liu, N., and Lai, E. C. (2009). Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes. Mol Cell *36*, 431-444.

Olivieri, D., Senti, K. A., Subramanian, S., Sachidanandam, R., and Brennecke, J. (2012). The Cochaperone Shutdown Defines a Grou*p of Biogen*esis Factors Essential for All piRNA Populations in *Drosophila*. Mol Cell *47*, 954-969.

Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K., and Aravin, A. A. (2013). Bacterial Argonaute samples the transcriptome to identify foreign DNA. Mol Cell *51*, 594-605.

Orban, T. I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. RNA *11*, 459-469.

Park, J. E., Heo, I., Tian, Y., Simanshu, D. K., Chang, H., Jee, D., Patel, D. J., and Kim, V. N. (2011). Dicer recognizes the 5' end of RNA for efficient and accurate processing. Nature *475*, 201-205.

Parker, J. S., Roe, S. M., and Barford, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. EMBO J *23*, 4727-4737.

Parker, J. S., Parizotto, E. A., Wang, M., Roe, S. M., and Barford, D. (2009). Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. Mol Cell *33*, 204-214.

Parker, J. S., Roe, S. M., and Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. Nature *434*, 663-666.

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Muller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P.,

Davidson, E., and Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature *408*, 86-89.

Pennisi, E. (2013). The CRISPR craze. Science *341*, 833-836.

Powell-Coffman, J. A., Knight, J., and Wood, W. B. (1996). Onset of C. elegans gastrulation is blocked by inhibition of embryonic transcription with an RNA polymerase antisense RNA. Dev Biol *178*, 472-483.

Preall, J. B., Czech, B., Guzzardo, P. M., Muerdter, F., and Hannon, G. J. (2012). shutdown is a component of the *Drosophila* piRNA biogenesis machinery. RNA *18*, 1446-1457.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173-1183.

Rajewsky, N., and Socci, N. D. (2004). Computational identification of microRNA targets. Dev Biol *267*, 529-535.

Rand, T. A., Petersen, S., Du, F., and Wang, X. (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell *123*, 621-629.

Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., and Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. RNA *11*, 1640-1647.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000). The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. Nature *403*, 901-96.

Reuter, J. S., and Mathews, D. H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics *11*, 129.

Reuter, M., Berninger, P., Chuma, S., Shah, H., Hosokawa, M., Funaya, C., Antony, C., Sachidanandam, R., and Pillai, R. S. (2011). Miwi catalysis is

required for piRNA amplification-independent LINE1 transposon silencing. Nature *480*, 264-267.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W. S., and Khvorova, A. (2004). Rational siRNA design for RNA interference. Nat Biotechnol *22*, 326-330.

Rice, P., and Mizuuchi, K. (1995). Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. Cell *82*, 209-220.

Rivas, F. V., Tolia, N. H., Song, J. J., Aragon, J. P., Liu, J., Hannon, G. J., and Joshua-Tor, L. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. Nat Struct Mol Biol *12*, 340-349.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res *42*, W320-W324.

Rocheleau, C. E., Downs, W. D., Lin, R., Wittmann, C., Bei, Y., Cha, Y. H., Ali, M., Priess, J. R., and Mello, C. C. (1997). Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. Cell *90*, 707-716.

Roovers, E. F., Rosenkranz, D., Mahdipour, M., Han, C. T., He, N., Chuva de Sousa Lopes, S. M., van der Westerlaken, L. A., Zischler, H., Butter, F., Roelen, B. A., and Ketting, R. F. (2015). Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Rep *10*, 2069-2082.

Ross, P. D., and Sturtevant, J. M. (1960). The Kinetics of Double Helix Formation from Polyriboadenylic Acid and Polyribouridylic Acid. Proc Natl Acad Sci U S A *46*, 1360-1365.

Ross, P. D., and Sturtevant, J. M. (1962). On the kinetics and mechanism of helix formation: The two stranded poly(A+U) complex from polyriboadenylic acid and polyribouridylic acid. Journal of the American Chemical Society *84*, 4503-4507.

Ruby, J. G., Stark, A., Johnston, W. K., Kellis, M., Bartel, D. P., and Lai, E. C.

(2007). Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. Genome Res *17*, 1850-1864.

Sabin, L. R., Delas, M. J., and Hannon, G. J. (2013). Dogma derailed: the many influences of RNA on the genome. Mol Cell *49*, 783-794.

Saito, K., Nishida, K. M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M. C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. Genes Dev *20*, 2214-2222.

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M. C. (2007). Pimet, the *Drosophila* homolog of HEN1, mediates *2′*-O-methylation of Piwi-interacting RNAs at their 3′ ends. Genes Dev *21*, 1603-1608.

Salomon, W., Bulock, K., Lapierre, J., Pavco, P., Woolf, T., and Kamens, J. (2010). Modified dsRNAs that are not processed by Dicer maintain potency and are incorporated into the RISC. Nucleic Acids Res *38*, 3771-3779.

Salomon, W. E., Jolly, S. M., Moore, M. J., Zamore, P. D., and Serebrov, V. (2015). Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. Cell *162*, 84-95.

Sasaki, H. M., and Tomari, Y. (2012). The true core of RNA silencing revealed. Nat Struct Mol Biol *19*, 657-660.

Schirle, N. T., Sheu-Gruttadauria, J., and MacRae, I. J. (2014). Structural basis for microRNA targeting. Science *346*, 608-613.

Schirle, N. T., Sheu-Gruttadauria, J., Chandradoss, S. D., Joo, C., and MacRae, I. J. (2015). Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets. eLife *4*,

Schirle, N. T., and MacRae, I. J. (2012). The crystal structure of human Argonaute2. Science *336*, 1037-1040.

Schrödinger, L. L. C. (2010). The PyMOL Molecular Graphics System, Version

1.3.

Schwarz, D. S., Ding, H., Kennington, L., Moore, J. T., Schelter, J., Burchard, J., Linsley, P. S., Aronin, N., Xu, Z., and Zamore, P. D. (2006). Designing siRNA that distinguish between genes that differ by a single nucleotide. PLoS Genet *2*, e140.

Schwarz, D. S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199-208.

Schwarz, D. S., Hutvagner, G., Haley, B., and Zamore, P. D. (2002). Evidence that siRNAs function as guides, not primers, in the *Drosophila* and human RNAi pathways. Mol Cell *10*, 537-548.

Schwarz, D. S., Tomari, Y., and Zamore, P. D. (2004). The RNA-Induced Silencing Complex Is a Mg^{2+}-Dependent Endonuclease. Curr Biol *14*, 787-791.

Segur, J. B., and Oberstar, E. (1951). Viscosity of Glycerol and Its Aqueous Solutions. Industrial and Engineering Chemistry *43*, 2117-2120.

Seitz, H., Tushir, J. S., and Zamore, P. D. (2011). A 5′-uridine amplifies miRNA/miRNA* asymmetry in *Drosophil*a by promoting RNA-induced silencing complex formation. Silence *2*, 1-10.

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature *455*, 58-63.

Semenova, E., Jore, M. M., Datsenko, K. A., Semenova, A., Westra, E. R., Wanner, B., van der Oost, J., Brouns, S. J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc Natl Acad Sci U S A *108*, 10098-10103.

Sheng, G., Zhao, H., Wang, J., Rao, Y., Tian, W., Swarts, D. C., van der Oost, J.,

Patel, D. J., and Wang, Y. (2014). Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. Proc Natl Acad Sci U S A *111*, 652-657.

Shin, C. (2008). Cleavage of the star strand facilitates assembly of some microRNAs into Ago2-containing silencing complexes in mammals. Mol Cells *26*, 308-313.

Shin, C., Nam, J. W., Farh, K. K., Chiang, H. R., Shkumatava, A., and Bartel, D. P. (2010). Expanding the microRNA targeting code: functional sites with centered pairing. Mol Cell *38*, 789-802.

Simon, B., Kirkpatrick, J. P., Eckhardt, S., Reuter, M., Rocha, E. A., Andrade-Navarro, M. A., Sehr, P., Pillai, R. S., and Carlomagno, T. (2011). Recognition of 2'-O-Methylated 3'-End of piRNA by the PAZ Domain of a Piwi Protein. Structure *19*, 172-180.

Song, J. J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., Hannon, G. J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat Struct Biol *10*, 1026-1032.

Song, J. J., Smith, S. K., Hannon, G. J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434-1437.

Song, R., Hennig, G. W., Wu, Q., Jose, C., Zheng, H., and Yan, W. (2011). Male germ cells express abundant endogenous siRNAs. Proc Natl Acad Sci U S A *108*, 13159-13164.

Sternberg, S. H., LaFrance, B., Kaplan, M., and Doudna, J. A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. Nature *527*, 110-113.

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., and Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9.

Nature *507*, 62-67.

Sugimoto, N., Kierzek, R., and Turner, D. H. (1987). Sequence dependence for the energetics of dangling ends and terminal base pairs in ribonucleic acid. Biochemistry *26*, 4554-4558.

Swarts, D. C., Jore, M. M., Westra, E. R., Zhu, Y., Janssen, J. H., Snijders, A. P., Wang, Y., Patel, D. J., Berenguer, J., Brouns, S. J., and van der Oost, J. (2014a). DNA-guided DNA interference by a prokaryotic Argonaute. Nature *507*, 258-261.

Swarts, D. C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R. F., Koonin, E. V., Patel, D. J., and van der Oost, J. (2014b). The evolutionary journey of Argonaute proteins. Nat Struct Mol Biol *21*, 743-753.

Tahbaz, N., Carmichael, J. B., and Hobman, T. C. (2001). GERp95 belongs to a family of signal-transducing proteins and requires Hsp90 activity for stability and Golgi localization. J Biol Chem *276*, 43294-43299.

Takimoto, K., Wakiyama, M., and Yokoyama, S. (2009). Mammalian GW182 contains multiple Argonaute-binding sites and functions in microRNA-mediated translational repression. RNA *15*, 1078-1089.

Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M., and Hannon, G. J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.

Tan, G. S., Garchow, B. G., Liu, X., Yeung, J., Morris, J. P., Cuellar, T. L., McManus, M. T., and Kiriakidou, M. (2009). Expanded RNA-binding activities of mammalian Argonaute 2. Nucleic Acids Res *37*, 7533-7545.

Tang, G., Reinhart, B. J., Bartel, D. P., and Zamore, P. D. (2003). A biochemical framework for RNA silencing in plants. Genes Dev *17*, 49-63.

Taylor, D. W., Zhu, Y., Staals, R. H., Kornfeld, J. E., Shinkai, A., van der Oost, J., Nogales, E., and Doudna, J. A. (2015). Structural biology. Structures of the

CRISPR-Cmr complex reveal mode of RNA target positioning. Science *348*, 581-585.

Teplova, M., Yuan, Y. R., Phan, A. T., Malinina, L., Ilin, S., Teplov, A., and Patel, D. J. (2006). Structural basis for recognition and sequestration of UUU(OH) 3' temini of nascent RNA polymerase III transcripts by La, a rheumatic disease autoantigen. Mol Cell *21*, 75-85.

Tian, Y., Simanshu, D. K., Ma, J. B., and Patel, D. J. (2011). Inaugural article: Structural basis for piRNA 2'-O-methylated 3'-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains. Proc Natl Acad Sci U S A *108*, 903-910.

Tolia, N. H., and Joshua-Tor, L. (2007). Slicer and the Argonautes. Nat Chem Biol *3*, 36-43.

Tomari, Y., Du, T., and Zamore, P. D. (2007). Sorting of *Drosophila* small silencing RNAs. Cell *130*, 299-308.

Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P. D. (2004). A protein sensor for siRNA asymmetry. Science *306*, 1377-1380.

Tomari, Y., and Zamore, P. D. (2005). Perspective: machines for RNAi. Genes Dev *19*, 517-529.

Turner, D. H., and Mathews, D. H. (2010). NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. Nucleic Acids Res *38*, D280-D282.

Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P. D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320-324.

van der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N. M., and Stuitji, A. R. (1990). Flavonoid genes in petunia: Addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell *2*, 291-299.

Voigt, F., Reuter, M., Kasaruho, A., Schulz, E. C., Pillai, R. S., and Barabas, O. (2012). Crystal structure of the primary piRNA biogenesis factor Zucchini

reveals similarity to the bacterial PLD endonuclease Nuc. RNA *18*, 2128-2134.

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. S., and Martienssen, R. A. (2002). Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. Science *297*, 1833-1837.

Wang, W., Yoshikawa, M., Han, B. W., Izumi, N., Tomari, Y., Weng, Z., and Zamore, P. D. (2014). The initial uridine of primary piRNAs does not create the tenth adenine that Is the hallmark of secondary piRNAs. Mol Cell *56*, 708-716.

Wang, X. H., Aliyari, R., Li, W. X., Li, H. W., Kim, K., Carthew, R., Atkinson, P., and Ding, S. W. (2006). RNA Interference Directs Innate Immunity Against Viruses in Adult *Drosophila*. Science *312*, 452-454.

Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T., and Patel, D. J. (2008a). Structure of an Argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature *456*, 921-926.

Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G. S., Tuschl, T., and Patel, D. J. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature *461*, 754-761.

Wang, Y., Sheng, G., Juranek, S., Tuschl, T., and Patel, D. J. (2008b). Structure of the guide-strand-containing Argonaute silencing complex. Nature *456*, 209-213.

Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev *20*, 1732-1743.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M. A.,

Sakaki, Y., and Sasaki, H. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

Wee, L. RNA Interference by the Numbers: Explaining Biology Through Ezymology. Worcester, MA: University of Massachusetts Medical School; 2013. p. Dissertation.

Wee, L., Flores-Jasso, C., Salomon, W., and Zamore, P. (2012). Argonaute Divides Its RNA Guide Into Domains with Distinct Functions and RNA-Binding Properties. Cell *151*, 1055-1067.

Wetmur, J. G. (1976). Hybridization and renaturation kinetics of nucleic acids. Annu Rev Biophys Bioeng *5*, 337-361.

Wetmur, J. G., and Davidson, N. (1968). Kinetics of renaturation of DNA. J Mol Biol *31*, 349-370.

Wiedenheft, B., van Duijn, E., Bultema, J. B., Waghmare, S. P., Zhou, K., Barendregt, A., Westphal, W., Heck, A. J., Boekema, E. J., Dickman, M. J., and Doudna, J. A. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc Natl Acad Sci U S A *108*, 10092-10097.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. Cell *75*, 855-62.

Xia, T., Mathews, D. H., and Turner, D. H. (2001). Thermodynamics of RNA Secondary Structure Formation. In RNA, Soll, D., S. Nishimura, and P. B. Moore, eds. (Oxford: Pergamon), pp. 21-48.

Xia, T., SantaLucia, J. J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry *37*, 14719-14735.

Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., Jacobsen, S. E., and Carrington, J. C. (2004). Genetic and Functional Diversification of Small RNA Pathways in Plants. PLoS Biol *2*, E104.

Xiol, J., Cora, E., Koglgruber, R., Chuma, S., Subramanian, S., Hosokawa, M., Reuter, M., Yang, Z., Berninger, P., Palencia, A., Benes, V., Penninger, J., Sachidanandam, R., and Pillai, R. S. (2012). A Role for Fkbp6 and the Chaperone Machinery in piRNA Amplification and Transposon Silencing. Mol Cell *47*, 970-979.

Yamashita, A., Chang, T. C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C. Y., and Shyu, A. B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat Struct Mol Biol *12*, 1054-1063.

Yan, K. S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M. M. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468-474.

Yang, D., Lu, H., and Erickson, J. W. (2000). Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in *Drosophila* embryos. Curr Biol *10*, 1191-1200.

Yang, N., and Kazazian, H. H. J. (2006). L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol *13*, 763-771.

Yang, W., and Steitz, T. A. (1995). Recombining the structures of HIV integrase, RuvC and RNase H. Structure *3*, 131-134.

Ye, X., Huang, N., Liu, Y., Paroo, Z., Huerta, C., Li, P., Chen, S., Liu, Q., and Zhang, H. (2011). Structure of C3PO and mechanism of human RISC activation. Nat Struct Mol Biol

Yi, R., Qin, Y., Macara, I. G., and Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev *17*,

3011-3016.

Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., and Tomari, Y. (2010). ATP-dependent human RISC assembly pathways. Nat Struct Mol Biol *17*, 17-23.

Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.

Zander, A., Holzmeister, P., Klose, D., Tinnefeld, P., and Grohmann, D. (2014). Single-molecule FRET supports the two-state model of Argonaute action. RNA Biol *11*, 45-56.

Zeng, Y., and Cullen, B. R. (2004). Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. Nucleic Acids Res *32*, 4776-4785.

Zha, X., Xia, Q., and Yuan, Y. A. (2012). Structural insights into small RNA sorting and mRNA target binding by Arabidopsis Argonaute Mid domains. FEBS Lett *586*, 3200-3207.

Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell *118*, 57-68.

Zhang, Z., Revyakin, A., Grimm, J. B., Lavis, L. D., and Tjian, R. (2014). Single-molecule tracking of the transcription cycle by sub-second RNA detection. eLife *3*, e01775.

Zheng, G., Ambros, V., and Li, W. H. (2010). Inhibiting miRNA in *Caenorhabditis elegans* using a potent and selective antisense reagent. Silence *1*, 9.